University of Tennessee, Knoxville
TRACE: Tennessee Research and Creative Exchange

Masters Theses | Graduate School

8-2019

Improving Identification Methods for Tabanus Flies (Diptera: Tabanidae) from the Southeastern United States using DNA Barcoding & Environmental Niche Modeling

Travis Davis
University of Tennessee, tdavi113@vols.utk.edu

Follow this and additional works at: https://trace.tennessee.edu/utk_gradthes

Recommended Citation

Davis, Travis, "Improving Identification Methods for Tabanus Flies (Diptera: Tabanidae) from the Southeastern United States using DNA Barcoding & Environmental Niche Modeling. " Master's Thesis, University of Tennessee, 2019.
https://trace.tennessee.edu/utk_gradthes/5501

This Thesis is brought to you for free and open access by the Graduate School at TRACE: Tennessee Research and Creative Exchange. It has been accepted for inclusion in Masters Theses by an authorized administrator of TRACE: Tennessee Research and Creative Exchange. For more information, please contact trace@utk.edu.

To the Graduate Council:

I am submitting herewith a thesis written by Travis Davis entitled "Improving Identification Methods for Tabanus Flies (Diptera: Tabanidae) from the Southeastern United States using DNA Barcoding & Environmental Niche Modeling." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Entomology and Plant Pathology.

Rebecca Trout Fryxell, Major Professor

We have read this thesis and recommend its acceptance:

John Moulton, Monica Papes

Accepted for the Council:
Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Improving Identification Methods for *Tabanus* Flies (Diptera: Tabanidae) from the Southeastern United States using DNA Barcoding and Environmental Niche Modeling

A Thesis Presented for the

Master of Science

Degree

The University of Tennessee, Knoxville

Travis Mitchell Davis

August 2019

Copyright © 2019 by Travis Mitchell Davis

All Rights Reserved

Acknowledgements

I would like to give a special thanks to my advisor Dr. Rebecca Trout Fryxell for her mentorship throughout all stages of this study. The authors of this study also offer a profound thanks the staff and personnel of the UTIA Research and Education Centers, and other collaborators (Muller, Nichols, Parys, Ray, Riles, Schiff, Stewart, Super, and Watson) for collecting horse flies for this study. This project was supported by the Department of Entomology and Plant Pathology and USDA-ARS.

Abstract

Blood-feeding female horse flies (Diptera: Tabanidae: *Tabanus*) are pests of livestock and man worldwide. Direct damage from *Tabanus* blood-feeding results in blood loss and physical damage to the skin. Indirect outcomes are the potential transmission of pathogens, economic losses in livestock production, and disruption of outdoor recreation. Horse flies are an understudied group and *Tabanus* classification remains incompletely resolved due to variable morphological characters, high diversity, and limited research within the group. Therefore, the first step to evaluating horse flies as pests is improving identification methods. Our overarching goal was to improve methods of *Tabanus* identification by building a DNA barcoding database of *Tabanus* flies and producing distribution models using environmental niche modeling (ENM) in the Southeastern U.S. To complete this objective, horse flies were collected with fly traps and opportunistically by researchers and collaborators throughout the study range. C*ytochrome oxidase subunit I* (COI) barcodes were sequenced for 40 horse fly species collected in 6 states in the Southeastern U.S. (Tennessee, Arkansas, Mississippi, Alabama, Florida, and North Carolina). Nine major clades were recovered with varying levels of posterior probability support that effectively help to identify *Tabanus* species. COI is proven to be an effective tool for identification and vector incrimination of most of the 40 *Tabanus* species represented. Environmental niche models (ENMs) were produced for the most pervasive horse flies collected in this study in the Southeastern U.S. (*T. fulvulus* Wiedemann*, T. lineola* Fabricius*, T. subsimilis* Bellardi*, T. quinquevittatus* Wiedemann*, T. sparus milleri* Whitney, and *T. sulcifrons* Macquart). The resulting distributions of these flies included locations with high relative humidity and temperature range (continentality) based on model analysis in 18 states throughout the Eastern U.S. Together, this work provides a reference point at the species level to further investigate biological differences among horse flies such as feeding behavior, host specificity, seasonality, and range to be evaluated in the determination of economic management options.

Table of Contents

List of Tables

List of Figures

List of Attachments

Introduction

The family Tabanidae (Diptera) contains more than 4,400 species worldwide (Pape et al., 2011). Much of the diversity is contained within the genus *Tabanus* and classification of this group remains incompletely resolved (Morita et al. 2016). In the Southeastern U.S. there are more than 68 species of horse flies (*Tabanus*) (Goodwin et al. 1985, Nalen et al. 2015). Immature larvae are aquatic to semi-aquatic predators and adults feed on the nectar of flowering plants as a source of energy, while only the female flies feed on the blood of vertebrates to nourish egg production (Philip S. Corbet 1962). Their persistent blood feeding behavior causes economic losses in livestock production due to loss of blood and animal weight, alteration of grazing behavior, and pathogen transmission (Foil and Hogsette 1994). In 1965, it was estimated that losses in beef cattle production attributable to Tabanids were around $40 million annually (Anon, 1979). In 1991 Losses attributable to horse fly damage on calves, stockers, and dairy operations in the United States were estimated at $190 million (Kunz et al. 1991). How these estimates translate to current loss estimates and how these losses may be accounted for with respect to horse fly species are crucial for any pest management decisions and yet to be known.

Members of *Tabanus* worldwide are also vectors of parasitic nematodes, bacteria, and viruses that cause disease in wildlife, livestock, and at times humans (Foil 1989). Disease causing parasites can be transmitted biologically or mechanically by horse flies. Biologically transmitted pathogens develop and multiply within the host prior to transmission. One such example is the parasitic nematode *Elaephora schneideri* which is transmitted to large ungulates (Couvillion et al. 1986, Hibler et al. 1971). Mechanical transmission is accomplished through contact with a contamination source, such as the blood of an infected host, and the transportation to another host via the body or mouthparts of the fly (Foil 1989). Biting is often interrupted due to the host's defensive responses, such as tail swiping and head throwing (Baldacchino et al. 2014). As a consequence, Tabanids may bite multiple hosts several times before obtaining, thus

contaminate healthy animals (Foil 1989). *Trypanosoma evansi*, which causes Surra, is mechanically transmitted to animals (and humans on rare occasion) throughout South America, Africa, and Asia (Brun et al. 1998). Surra is one of the most widespread trypanosome parasites and of particular concern in regions such as Spain, Australia, and France where it has shown signs of becoming an emerging disease (Desquesnes et al. 2013). Most livestock are susceptible to *Trypanosoma evansi* and once Infected, hosts may experience miscarriage, acute fever, lethargy, anemia, and death (Gill, 1977). Throughout the world equine infectious anemia virus (EIA), *Anaplasma marginale*, and *Bacillus anthraxis* are among pathogens mechanically transmitted by horse flies (N.S. Krishna Rao 1958, Tidwell et al. 1972, Hawkins et al. 1976, 1982). EIA disease is a potentially fatal and very costly viral disease that has remained a problem in the United States for decades (Krinsky 1976). Figure 1 illustrates the most recent EIA statistics for 2016 in the U.S published by the U.S. Department of Agriculture Animal and Plant Health Inspection Service ("Equine Infectious Anemia Distribution Maps" 2017). There is no vaccine for EIA and infected individuals are carriers for life. Animals that test positive for the virus are typically euthanized or quarantined for life (Pérez et al. 2002).

Management techniques and methods for Tabanids in pastured livestock include avoidance of high density areas, the use of traps to intercept blood-seeking females, providing shelter for livestock, and use of pesticides; however, these techniques have not effectively addressed all the problems associated with losses in livestock productivity and losses in recreational activities(Axtell 1976, Foil 1989, Foil and Hogsette 1994). This is notably due to the lack of information available on the life cycles and roles of important pest species and limited capability to measure the direct impact of Tabanidae within their respective niches. Improving Tabanid identification and distributional knowledge through molecular and niche modeling tools is the first crucial step to gathering needed biological data that will help clarify the economic impact and ecology of horse flies (genus: *Tabanus*) in the Southeastern U.S. (Meier et al. 2006, Costa

and Carvalho 2010, Cywinska et al. 2010, Banerjee et al. 2015). Cytochrome oxidase subunit 1 (COI) will provide basic evolutionary history information for species identification but may not be sufficient for identification of every *Tabanus* fly to species. Another very useful tool for developing more effective pest management techniques in generating distribution maps using maximum entropy environmental niche modeling. Ranges of species can be estimated using Maxent software to map the most likely distribution of a given species.

Historically, *Tabanus* classification has been difficult due to ambiguous morphological characters, high variation and group diversity, and the limited availability of experts and research within the group (Kristen Bartlett, Steven R. ALM, Roger Lebrun 2002, Lessard and Yeates 2011, Banerjee et al. 2015). Chaetotaxy, commonly used among Diptera is not useful for *Tabanus* flies (Teskey and Pechuman 1983). Recently, molecular systematic tools using mitochondrial (COI), mitochondrial 16s rRNA, and nuclear (CAD, AATS) genes has shown promise in revising and breathing new life into the family (Morita et al. 2016). The same study found the genus *Tabanus* to be a non-monophyletic group within the family Tabanidae due to the current classification based predominately on morphological identification (Morita et al. 2016). The lack of stable morphological characters to depend on for identification rationalizes the use of molecular tools and distribution modeling for identification. The use of molecular tools is more accurate and dependable and environmental niche modeling to improves ecological knowledge of identified species; together, these tools support species distinctions at a molecular level, and provide good foundations for future ecological and biological studies, which are the intent of the objectives of this study.

Chapter 1

Improving the identification of *Tabanus* flies through the development of COI DNA barcodes

Abstract

Blood-feeding horse flies cause economic and health problems worldwide, but management of these pestiferous flies remains ineffective due to difficulties in horse fly identification and research. *Tabanus* is a specious and non-monophyletic clade worldwide and classification remains incompletely resolved due to unstable morphological characters, high diversity, cryptic species, and limited expertise and research within the group. The concerns with morphological identification of *Tabanus* have stifled the progress of an already understudied group. To facilitate research efforts aimed at biological studies evaluating *Tabanus* as pests in the Southeastern U.S., *cytochrome oxidase subunit I* (COI) fragments were sequenced for 40 horse fly species collected in 6 states in the Southeastern U.S. (Tennessee, Arkansas, Mississippi, Alabama, Florida, and North Carolina) to bolster identification. Bayesian analysis of sequences derived from these 40 *Tabanus* morphospecies resulted in species-specific clades having moderate to high posterior probability values (0.95 - 1.0). Nine major clades were recovered that effectively help to identify Tabanus species. Individuals within species of presumed or established complexes and those having similar morphologies are recovered separately with high posterior probability in the majority but not all cases. COI is an effective tool for identification and vector incrimination of most of the 40 *Tabanus* species represented and contributes to the goal of resolving classification within *Tabanus*.

Introduction

The lacerating, blood-feeding behavior of female horse flies causes stress to animal health and livestock production systems worldwide. The direct damage caused by potentially hundreds of flies biting and feeding on an animal occurs through blood loss, behavioral disturbance, and stress (Foil and Hogsette 1994). Host responses to this irritation are observed in defensive behaviors such as tail flicks, head tosses, and stomping (Mooring et al. 2007, Baldacchino et al. 2014). All categories of damage caused by horse flies in livestock production systems are

attributed to and measured in average losses in weight gain in comparison to protected animals such as those treated with an insecticide (Perich et al. 1986). Horse flies also transmit pathogens either biologically (*Elaephora schneider*) or mechanically (*Anaplasmosis marginale*), and host's hides are often damaged from skin lacerations, which is significant in beef cattle operations (Perich et al. 1986). Horse flies can also cause nuisance and harm to humans and have the potential to illicit allergic reactions, cause infection, and make outdoor recreation in infested areas insufferable (Goodwin and Bastiaan 1996, Quercia et al. 2008).

It is estimated that as many as 68 species of *Tabanus* occurring in the Southeastern U.S. east of the Mississippi river, based on previous collection records (Jones and Anthony 1964, Goodwin et al. 1985, Nalen et al. 2015). The southern extremes of this range including the southern Gulf and Atlantic coasts and southern Florida contain 18 species (*T. acutus* Bigot*, T. birdiei* Whitney*, T. cayensis* Fairchild*, T. cheliopterus* Rondani*, T. coarctatus* Stone*, T. colon* Thunberg*, T. conterminous* Walker*, T. daedalus* Stone, *T. endymion* Osten Sacken*, T. fronto* Osten Sacken*, T. fulvilineis* Philip*, T. fumipennis* Wiedemann*, T. fusconervus* Macquart*, T. hinellus* Philip*, T. kisliuki* Stone*, T. triunctus* Walker*, T. vittiger* Hines*, T. yucatanus* Townsend) that are seemingly restricted in range and not collected in this study. Explanations for the narrow known distributions of this southern subset of *Tabanus* diversity are speculated to be due to collection bias or potentially the introduction of Caribbean species that have become established (Nalen et al. 2015). In the state of Tennessee, over 20 species of horse flies were cataloged as pests of humans and livestock (Goodwin et al. 1985); however, the degree to which economic loses are sustained from each *Tabanus* species and in what capacity is yet to be quantified. Information on feeding behavior and ecology exists for several species in the southeastern U.S. and is generalized for much of the remaining diversity (Foil 1989, Friend and Stoffolano 1991). Owing to this deficit in information are the difficulties inherent with morphological identification of horse flies.

The difficulties of horse fly identification are a result of their unstable morphological characters, the high diversity of and presence of cryptic species within *Tabanus*, limited research, and loss of taxonomic specialists (Teskey 1969, Bartlett et al. 2002, Lessard and Yeates 2011, Banerjee et al. 2015). Closely related species groups with similar morphology, ambiguous genital characters, variation in dorsal patterning, antennal, and frons characters require more scrutiny to address species complexes and corroborate current classification with biological data within *Tabanus*. Several *Tabanus* species complexes occur within the Southeastern U.S., such as the *T. lineola* complex, which is thought to include at least 20 species throughout the Americas (Goodwin and Drees 1996).

Male and female horse flies are sexually dimorphic and separate keys are required for species identification. *Tabanus* males are rarely captured or studied. Immature stages are not well studied apart from morphological descriptions and occurrence (Teskey and Pechuman 1983, Goodwin et al. 1985). Even with the additional collection of male and immature horse flies, their morphological identification is presumed to be more challenging than females. This is because male horse flies lack the calli structure between the female's dichoptic eyes and have fewer distinguishing characters. Larval identifications are dependent on pale pubescence that is difficult to observe (Teskey and Pechuman 1983, Goodwin and Bastiaan 1996). With challenging prospects for morphological identification of all life stages, it can be hypothesized that the use of DNA barcoding will improve identification of *Tabanus* species.

Within Diptera, the COI gene was found to have an average sequence divergence of 9%; however, average congeneric COI sequence divergence rates vary significantly by geographic region within family groups and require analysis before any other utility can be determined (Hebert et al. 2003, Cywinska et al. 2010). Knowing this, Cywinska et al. (2010) evaluated its use as a method to identify species and reported that it worked with 4 *Tabanus* species (*T.*

atratus Fabricius, *T. marginalis* Wiedmann, *T. rufofrater* Walker, *T. similis* Macquart) and had a 5.96% divergence rate between the western Canadian collections. Later in India, Banjeree et al. (2015) used the COI gene to identify *T.* striatus and eventually incriminate it as the local horse fly vector of the protozoan causing Surra in livestock. Based on these results, it appears that COI is a useful marker that can be used as a practical tool for taxonomic identification of tabanids at the species level. Here, my objective is to improve methods of *Tabanus* identification by building a COI barcoding database of *Tabanus* species by testing the hypothesis that we will be able to aid horse fly identification with unique barcode sequences that are distinct to identify the different and diverse horse flies present in the Southeastern U.S.

Materials and Methods

Tabanus collection and identification

Material for this study came from a combination of sources including actively collecting *Tabanus* flies with traps and through submissions from collaborators throughout the Southeastern U.S. All specimens were collected from 2014-2018 in Alabama (1 site), Arkansas (1 site), Florida (14 sites), Mississippi (14 sites), North Carolina (2 sites), and Tennessee (42 sites). Trapping sites were selected among appropriate locations with suitable tabanid habitat and accessibility to maintain traps. Traps were installed in locations near animal enclosures having direct sunlight and were physically separated by at least 1 kilometer where possible.

A number of different methods were used to collect *Tabanus* flies, including incidental collections, but three traps were primarily used: The H trap (Bite-Lite, Bethel, CT), the NZI trap (Rincon-Vitova, Ventura, CA), and the I trap. I designed the I trap which features an aluminum cage structure cut and formed from Garden Zone gray Steel Hardware cloth approximately 4ft x 4ft funnel shaped from hardware cloth, wrapped in a phthalogen blue tarp, and suspended from

an 8ft garden hook. Materials and supplies used for construction included a section of Garden Zone's 6ftx2ft length hardware cloth (Lowe's), Blue Hawk 14-gauge Multipurpose wire (Lowe's), Saint-Gobain ADFORS 3ft x 25ft Gray Aluminum Replacement Screen (Lowe's), a Polypropylene Powder Funnel (180mm) (Globe Scientific, Mahwah, NJ), a 20oz and a 32oz plastic bottle, a hopper ball (AppleRound, Phoenix, AZ) spray-painted black, and plastic bottles ranging from 0.5-1.0L for collection heads. The assembly protocol can be found in Appendix 1.

At active trapping sites, collection heads were removed and replenished with 300ml of 80% ethanol twice a week from May to October. Collected material was placed into a 50mL falcon tube (ThermoFisher Scientific, Hampton, NH) containing 80% ethanol, labeled by trap, location, and date, and then stored in a -20°C freezer until identified. Captured tabanid flies were first identified to genus and then *Tabanus* spp. were identified to species using descriptive dichotomous? keys (Drees et al. 1980, Goodwin et al. 1985). Additional resources used to confirm taxonomic identification included specimens in the following repositories: the University of Tennessee Entomology and Plant Pathology Insect Museum, the Cornell Insect Collection, the University of Auburn Insect Collection, and the Florida State Collection of Arthropods.

Specimen selection and COI amplification

Following the identification of *Tabanus* flies to species, carefully selected specimens were identified as voucher specimens for both museum and genetic archiving (Table 1). Specimens were selected based on their preservation quality, geographic origin, and morphological condition. Identified specimens were confirmed by Dr. James T. Goodwin. In total, 215 specimens representing 40 species were selected for amplification of the c*ytochrome oxidase subunit I* (COI) gene, often referred to as DNA barcoding gene. *Chrysops moechus* collected from Little River Dairy Farm was sequenced as an outgroup to *Tabanus*. Previous published

projects (Deng and Hiruki 1991, Folmer et al. 1994) and unpublished projects in the Trout Fryxell laboratory used the COI gene to distinguish different *Tabanus* species. An approximately 600 bp region within the COI gene is used as a species identification tool for many animal groups including arthropods (Hebert et al. 2003) such as *Tabanus* (Cywinska et al. 2010, Banerjee et al. 2015).

To preserve the specimen and generate the physical voucher specimen, DNA was extracted from the three right legs of each selected specimen using sterile technique. Dissected legs were initially digested in a solution of proteinase K (20uL) and buffer ATL (180uL). A 5mm tungsten carbide bead (5mm) was added to each extraction tube and homogenized in a TissueLyser II (Qiagen, Hilden, Germany) at 15Hz for 20s. Samples were incubated in a Max 4450 shaking incubator (ThermoFisher Scientific) at 56°C in the lysis solution overnight. Automated high-throughput DNA extraction was performed using the HT system (Qiagen) for up to 96 samples and eluted at 200 microliters in AE per sample.

Individual polymerase chain reactions (PCRs) were performed in 25µL reactions comprised of 3µl of extracted DNA, 10 µl of Maxima Hotstart PCR mix (ThermoFisher Scientific), 1µl of each forward (5'-GGTCAACAAATCATAAAGATATTGG) and reverse (5'-TAAACTTCAGGGTGA CCAAAAAATCA) primer, and 10 µl of PCR-grade water (ThermoFisher Scientific). Reaction mixes were performed under a PCR Enclosure hood (Labconco, Kansas City, MO). Thermocycling conditions consisted of a 1 min initial denaturing at 95 ºC, 35 cycles of denaturing at 94 ºC for 1 min, annealing at 55ºC for 1 min, and a final extension for 1.5 min at 72 ºC as previously described (Folmer et al. 1994). Samples were held at 4ºC in the thermocycler until they were stored at -20ºC for gel electrophoresis. Table-thawed amplified reactions were visualized in a 0.5% agarose gels consisting of 4.5g of Agarose (ThermoFisher

Table 1. *Tabanus* specimen vouchers. Bracketed are the number of specimens, and the number of localities per species. Sex and location of collection site are also listed.

Species	Voucher Specimen Identification in Tennessee Museum	Sex (M/F)	Collection Location County, State
Tabanus atratus (7,7)	18tab-HLRM2-Tatr-0086 18tab-PLAT3-Tatr-0044 18tab-SF2-Tatr-0055 18Tab-AMES5-Tatr-14373 18tab-AMES7-Tatr-14673 18tab-HLRM3-Tatr-0034 18Tab-LRIV3-Tatr-0076	F F F F M F F	Robertson, TN Cumberland, TN Bay, FL Fayette, TN Fayette, TN Robertson, TN Blount, TN
Tabanus abdominalis (5,1)	18tab-AMES6-Tabd-9642 18tab-AMES6-Tabd-16998 18tab-AMES6-Tabd-15898 18tab-AMES6-Tabd-15899 18tab-AMES6-Tabd-16999	F F F F F	Fayette, TN Fayette, TN Fayette, TN Fayette, TN Fayette, TN
Tabanus americanus (5,4)	18tab-AMES6-Tamer-9336 18tab-AMES6-Tamer-9589 18tab-AMES7-Tamer-6225 18tab-MSWA13-Tamer-0419 18tab-MSWA5-Tamer-0472	F F F F F	Fayette, TN Fayette, TN Fayette, TN Sunflower, MS Sunflower, MS
Tabanus aranti (4,2)	18tab-AMES6-Taranti-10534 18tab-AMES6-Taranti-8856 18tab-AMES7-Taranti-6312 18tab-AMES7-Taranti-6951	F F F F	Fayette, TN Fayette, TN Fayette, TN Fayette, TN
Tabanus calens (6,6)	18tab-HOLS1-Tcal-0071 18tab-MSWA2-Tcal-0052 17tab-DBWR-Tcal-0009 18tab-AMES6-Tcal-15896 18tab-AMES7-Tcal-14972 18tab-LRIV2-Tcal-0671	F F F F F F	Knox, TN Tallahatchie, MS Arkansas, AR Fayette, TN Fayette, TN Blount, TN
Tabanus cymatomorphous (1,1)	18tab-MSWA13-Tcym-0420	F	Bolivar, MS
Tabanus equalis (3,2)	18tab-MSWA7-Tequal-0390 18tab-MSWA7-Tequal-0391 18tab-MSWA9-Tequal-0393	F F F	Bolivar, MS Bolivar, MS Bolivar, MS

Table 1 (continued)

Species	Voucher Specimen Identification in Tennessee Museum	Sex (M/F)	Collection Location County, State
Tabanus fairchildi (5,2)	18Tab-AU1-Tfair-0004 18tab-LRIV3-Tfair-0474 18tab-LRIV3-Tfair-0607 18tab-LRIV3-Tfair-0543 18tab-LRIV3-Tfair-0544	F F F F F	Lee, AL Blount, TN Blount, TN Blount, TN Blount, TN
Tabanus fulvulvus (6,6)	18tab-LRIV3-Tfulv-0137 18tab-Ames6-Tfulv-1953 18tab-GRASS2-Tfulv-0012 18tab-HOLS3-Tfulv-0034 18tab-NC1-Tfulv-0007 18tab-MSWA5-Tfulv-0448	F F F F F F	Blount, TN Fayette, TN Cumberland, TN Knox, TN Lee, NC Sunflower, MS
Tabanus gladiator (4,3)	18tab-AMES7-Tglad-14822 18Tab-MSWA8-Tglad-0392 18Tab-SF2-Tglad-0056 18Tab-SF2-Tglad-0057	F F F F	Fayette, TN Wilkinson, MS Bay, FL Bay, FL
Tabanus imitans (1,1)	18tab-AMES7-Timi-6229	F	Fayette, TN
Tabanus limbatinevris (5,5)	18tab-HLRM2-Tlimb-0069 18tab-MSWA5-Tlimb-0468 18tab-AMES4-Tlimb-14003 18tab-AMES6-Tlimb-16997 18tab-AMES7-Tlimb-9244	F F F F F	Robertson, TN Sunflower, MS Fayette, TN Fayette, TN Fayette, TN
Tabanus lineola (7,7)	18tab-NC1-Tlin-0008 18Tab-SF1-Tlin-0037 18Tab-SF2-Tlin-0052 16tab-MSWA1-Tlin-0303 18tab-AMES7-Tlin-10743 18Tab-AU1-Tlin-0002 18tab-Noxu1-Tlin-0001	F F F F F F F	Lee, NC Bay, FL Bay, TN Tallahatchie, MS Fayette, TN Lee, AL Noxubee, AL
Tabanus longiusculus (3,3)	18tab-AMES7-Tlongi-14672 18tab-PLAT1-Tlongi00026 18tab-AMES6-Tlongi-16382	F F F	Fayette, TN Cumberland, TN Fayette, TN

Table 1 (continued)

Species	Voucher Specimen Identification in Tennessee Museum	Sex (M/F)	Collection Location County, State
Tabanus longus (6,6)	18tab-LRIV0-Tlong-0073 18tab-AMES6-Tlong-16320 18tab-PLAT3-Tlong-0053 18tab-GSMNP2-Tlong-0009 18tab-LRIV0-Tlong-0075 18tab-GSMNP1-Tlong-0003	F F F F F F	Blount, TN Fayette, TN Cumberland, TN Haywood, NC Blount, TN Sevier, TN
Tabanus melanocerus (6,6)	18tab-AMES6-Tmelan-10346 18tab-GRASS2-Tmelan-0179 18tab-GSMNP1-Tmelan-0007 18tab-SF2-Tmelan-0064 18tab-Knox2-Tmelan-0013 18tab-LRIV3-Tmelan-0470	F F F F F F	Fayette, TN Cumberland, TN Sevier, TN Bay, TN Knox, TN Blount, TN
Tabanus moderator (6,3)	18tab-AMES5-Tmod-2150 18tab-AMES7-Tmod-10609 18tab-LRIV3-Tmod-0391 18tab-LRIV3-Tmod-0388 18tab-LRIV3-Tmod-0389 18tab-LRIV3-Tmod-0390	F F F F F F	Fayette, TN Fayette, TN Blount, TN Blount, TN Blount, TN Blount, TN
Tabanus mixis (6,6)	18tab-HOLS1-Tmmix-0065 18tab-Knox2-Tmmix-0003 18tab-AMES5-Tmmix-3035 18tab-HLRM3-Tmmix-0016 18tab-Knox1-Tmmix-0002 18tab-LRIV2-Tmmix-0303	F F F F F F	Knox, TN Knox, TN Fayette, TN Robertson, TN Knox, TN Blount, TN
Tabanus molestus (7,7)	18tab-AMES6-Tmmol-10554 18tab-GRASS2-Tmmol-0024 18tab-GSMNP1-Tmmol-0001 18tab-Knox1-Tmmol-0010 18tab-AMES7-Tmmol-6230 18tab-LRIV2-Tmmol-0313 18tab-NC1-Tmmol-0042	F F F F F F F	Fayette, TN Cumberland, TN Sevier, TN Knox, TN Fayette, TN Blount, TN Fayette, TN
Tabanus mularis (4,4)	18tab-AMES5-Tmul-14449 18tab-AMES7-Tmul-10226 18tab-MDTN1-Tmul-0023 18tab-MDTN3-Tmul-0028	F F F F	Fayette, TN Fayette, TN Maury, TN Maury, TN

Table 1 (continued)

Species	Voucher Specimen Identification in Tennessee Museum	Sex (M/F)	Collection Location County, State
Tabanus nigrescens (3,2)	18tab-AMES6-Tnigrisc-9381 18tab-AMES6-Tnigrisc-9382 18tab-AMES7-Tnigrisc-9203	F F F	Fayette, TN Fayette, TN Fayette, TN
Tabanus nigripes (5,5)	18Tab-AMES4-Tnigrip-13964 18tab-AMES6-Tnigrip-15446 18tab-LRIV3-Tnigr-0597 18Tab-SF1-Tnigrip-0039 18Tab-SF2-Tnigrip-0051	F F F F F	Fayette, TN Fayette, TN Blount, TN Bay, FL Bay, FL
Tabanus nigrovittatus (2,2)	18Tab-SF1-Tnigrov-0071 14tab-Bien-Tnigrov-0371	F F	Fayette, TN Scott, MS
Tabanus pallidescens (5,5)	18tab-GRASS2-Tpalli-0010 18tab-Knox1-Tpalli-0005 18tab-LRIV3-Tpalli-0221 18tab-MDTN3-Tpalli-0014 18tab-AMES7-Tpalli-14252	F F F F	Cumberland, TN Knox, TN Blount, TN Maury, TN Fayette, TN
Tabanus petiolatus (5,5)	18tab-AMES7-Tpet-10680 18tab-LRIV3-Tpet-0699 18tab-NC1-Tpet-0006 18tab-AMES4-Tpet-14020 18Tab-SF1-Tpet-0040	F F F F F	Fayette, TN Blount, TN Haywood, NC Fayette, TN Bay, FL
Tabanus proximus (7,7)	18tab-AMES7-Tprox-14618 18tab-HLRM2-Tprox-0068 18tab-Knox1-Tprox-0016 18tab-LRIV3-Tprox-0696 18tab-MSWA11-Tprox-0401 18tab-MSWA13-Tprox-0437 18tab-NC1-Tprox-0044	F F F F F F F	Fayette, TN Robertson, TN Knox, TN Blount, TN Bolivar, MS Bolivar, MS Haywood, NC
Tabanus pumilus (3,3)	18tab-AMES5-Tpumi-0004 18tab-AMES7-Tpumi-3953 18tab-LRIV3-Tpumi-0145	F F F	Fayette, TN Fayette, TN Blount, TN

Table 1 (continued)

Species	Voucher Specimen Identification in Tennessee Museum	Sex (M/F)	Collection Location County, State
Tabanus quinquevittatus (7,7)	18tab-HLRM3-Tquin-0065 18tab-HOLS3-Tquin-0066 18tab-LRIV3-Tquin-0608 18tab-MDTN1-Tquin-0032 18tab-PLAT1-Tquin-0037 18tab-AMES6-Tquin-15227 17tab-LRIV0-Tquin-0007	F F F F F F F	Robertson, TN Knox, TN Blount, TN Maury, TN Cumberland, TN Fayette, TN Blount, TN
Tabanus reinwardtii (2,2)	18tab-AMES7-Trein-8474 18tab-LRIV3-Trein-0429	F F	Fayette, TN Blount, TN
Tabanus rufofrater (2,1)	17tab-ALACH5-Trufo-0035 17tab-ALACH5-Trufo-0036	F F	Alachua, FL Alachua, FL
Tabanus sackeni (6,6)	18tab-NC1-Tsack-0046 18tab-AMES5-Tsack-14399 18tab-AMES6-Tsack-9654 18tab-AMES7-Tsack-14610 18tab-GRASS1-Tsack-0169 18tab-GRASS3-Tsack-0061	F F F F F F	Haywood, NC Fayette, TN Fayette, TN Fayette, TN Cumberland, TN Cumberland, TN
Tabanus sparus milleri (7,7)	18tab-Knox2-Tsparm-0004 18tab-AMES5-Tsparm-3057 18tab-AMES7-Tsparm-3548 18tab-GRASS2-Tsparm-0139 18tab-HLRM2-Tsparm-0040 18tab-LRIV3-Tsparm-0430 18tab-NC1-Tsparm-0028	F F F F F F F	Knox, TN Fayette, TN Fayette, TN Cumberland, TN Robertson, TN Blount, TN Haywood, NC
Tabanus stygius (4,3)	17tab-DBWR-Tstyg-0001 18tab-AMES7-Tstyg-6228 18tab-MSWA5-Tstyg-0471 18tab-AMES7-Tstyg-5239	F F F F	Arkansas, AR Fayette, TN Sunflower, MS Fayette, TN
Tabanus sublongus (5,3)	18tab-LRIV3-Tsublong-0715 17tab-LRIV0-Tsublong-0071 18tab-AMES6-Tsublong-16318 17tab-LRIV0-Tsublong-0018 17tab-LRIV0-Tsublong-0038	F F F F F	Blount, TN Blount, TN Fayette, TN Blount, TN Blount, TN

Table 1 (continued)

Species	Voucher Specimen Identification in Tennessee Museum	Sex (M/F)	Collection Location County, State
Tabanus subsimilis (11,11)	18Tab-GRASS1-Tsubsim-0020	F	Cumberland, TN
	18Tab-MSWA10-Tsubsim-0395	F	Montgomery, MS
	18Tab-MSWA7-Tsubsim-0399	F	Bolivar, MS
	17tab-DBWR-Tsubsim-0003	F	Arkansas, AR
	18tab-HLRM3-Tsubsim-0005	F	Robertson, TN
	18Tab-LRIV2-Tsubsim-0681	F	Blount, TN
	18tab-MDTN3-Tsubsim-0040	F	Maury, TN
	18Tab-MSWA12-Tsubsim-0436	F	Bolivar, MS
	18Tab-MSWA13-Tsubsim-0438	F	Bolivar, MS
	16tab-MSWA1-Tsubsim-0513	F	Tallahatchie, MS
	16tab-LRIV0-Tsubsim-0004	F	Blount, TN
Tabanus sulcifrons (18,11)	18tab-MSWA13-Tsulc-0417	F	Bolivar, MS
	18tab-AMES7-Tsulc-15169	M	Fayette, TN
	18tab-DBWR-Tsulc-0005	F	Arkansas, AR
	18tab-Knox5-Tsulc-0015	F	Knox, TN
	18tab-MSWA12-Tsulc-0421	F	Bolivar, MS
	18tab-NC1-Tsulc-0045	F	Haywood, NC
	18tab-Scott1-Tsulc-0001	F	Scott, TN
	16-LRriv-LTsulc-0211	F	Blount, TN
	16-LRiv-LTsulc-0208	F	Blount, TN
	16-LRiv-LTsulc-0212	F	Blount, TN
	15-Sunfower-Etsulc-0175	F	Sunflower, TN
	16-LRiv-LTsulc-0206	F	Blount, TN
	17-Chatham Ltsulc 0185	F	Chatham, TN
	17-Chatham Ltsulc 0179	F	Chatham, TN
	17tab-MADI-Tsulc-0004	F	Madison, TN
	17tab-MADI-Tsulc-0005	F	Madison, TN
	17tab-MADI-Tsulc-0006	F	Madison, TN
	17tab-MADI-Tsulc-0007	F	Madison, TN
Tabanus trimaculatus (7,7)	18tab-HOLS1-Ttri-0037	F	Fayette, TN
	18tab-LRIV3-Ttri-0469	F	Blount, TN
	18tab-MDTN3-Ttri-0001	F	Maury, TN
	18tab-PLAT2-Ttri-0015	F	Cumberland, TN
	18tab-AMES6-Ttri-15326	F	Fayette, TN
	18tab-HLRM2-Ttri-0011	F	Robertson, TN
	18tab-NC1-Ttri-0029	F	Haywood, NC

Table 1 (continued)

Species	Voucher Specimen Identification in Tennessee Museum	Sex (M/F)	Collection Location County, State
Tabanus turbidus (5,5)	18tab-AMES5-Tturb-10856 18tab-AMES7-Tturb-10108 18tab-LRIV3-Tturb-0217 18tab-AMES6-Tturb-10487 18Tab-AuburnU1-Tturb-0001	F F F F F	Fayette, TN Fayette, TN Blount, TN Fayette, TN Lee, AL
Tabanus venustus (1,1)	15Tab-MSWA5-Tvenu-0017	F	Sunflower, MS

Scientific), 300mL of 1xTAE buffer diluted from 50x TAE buffer (ThermoFisher Scientific) with nuclease-free water, and 10µL of ethidium bromide). Agarose gels were loaded with 5uL of PCR product ran in a Horizontal Electrophoresis System (ThermoFisher Scientific) for 60min at 100V. Previously extracted *Tabanus* DNA was used as a positive control, negative controls consisted of PCR mastermix, and 4 lanes were dedicated to the 100bp DNA ladder (ThermoFisher Scientific). Samples producing a 600 bp amplicon were purified for DNA sequencing using ExoSAP-IT™ PCR Product Cleanup Reagent (ThermoFisher Scientific). Cleaned amplified product was then bidirectionally sequenced with Sanger sequencing at the University of Tennessee DNA Sequencing Facility on an Applied Biosystems 3730 Genetic Analyzer (ThermoFisher Scientific).

COI sequence analysis

The bidirectional raw sequences were aligned and reconciled using Bioedit Sequence Alignment Editor (Hall 1999). Consensus COI sequences were created for each species, representing between 1 and 18 specimens. These specimen sequences were then aligned using ClustalW in Bioedit, and species consensus species COI sequences were created. These resulting consensus species COI sequences were analyzed with Bayesian inference and coalescent theory.

To reconstruct the *Tabanus* species phylogenetic tree a strict molecular clock was implemented because sequence divergence among samples is relatively low (<10%) (Nascimento et al. 2017). Coalescent theory priors transition transversion ration (kappa), Bayes factor partition, root height of tree partition, and partition of coalescent population size were used to fit the model. 1,000,000 MCMC iterations were run in BEAST 1.10.4 and phylogenetic trees were visualized on FigTree v1.4.4 (Drummond et al. 2012, Suchard et al. 2018). Final phylogenetic

tree figures were edited to enhance legibility in Canvas™ 8 (https://www.canvasgfx.com/en/products/canvas-x-gis-2019/).

Results

In the phylogeny (Figure 1), 275 COI sequences were generated from 206 *Tabanus* specimens of 40 *Tabanus* species and 1 *Chrysops moechus* outgroup. *Tabanus nigrovittatus*, *T. superjumentarius*, *T. imitans,* and *T. cymatomorphous* were only collected once and represented in the phylogeny by one sequence. The phylogeny contains 4547 unique clades depicted with the cladogram and the phylogram with 9 major clades (A-I) recovered. Each clade is displayed (Figures 2-10) with the species and sequence labels listed for each taxon. Metadata of these taxa are given in Table 1 which includes the sex, number of specimens, and collection locations used to generate the respective COI sequences.
Posterior probability values represent the probability of each clades' monophyly (correctness) given the data.

A probability value greater than or equal to 0.95 is considered highly supported (Nascimento et al. 2017). Posterior probability values greater than or equal to 0.90 are represented in Figures 1-10. In Figure 1, lineages A-I are distinguished by high posterior support ($n \geq 0.90$) and/or distinct morphologies. Posterior probability values are only listed for clades with support greater than 0.90. Clade A (Figure 2) is described as flies with similar morphology and median dorsal triangles and is sister to Clade B, the commonly named greenhead flies (Figure 3), with high support (0.962). Clade C (Figure 4) contains *T. sparus milleri* and is weakly supported as the sister to the most recent common ancestor of clades D-I. Clade D (Figure 5) represents the lineage of the *lineola complex* and is moderately supported (0.9171) as sister to the ancestor to clades E-I. Clade E (Figure 6) contains *T. pumilus* and is weakly supported as sister to the most

Figure 1. COI Phylogeny. 1. Represents the cladogram and 2. Represents the phylogram of the *Tabanus* phylogeny of COI gene sequences 600bp in length. 206 taxa including 40 *Tabanus* species and *Chrysops moechus* outgroup are represented. Posterior probability values above 0.90 are listed for the clades labeled A-I. Pictures of horse fly species are listed corresponding to clades.

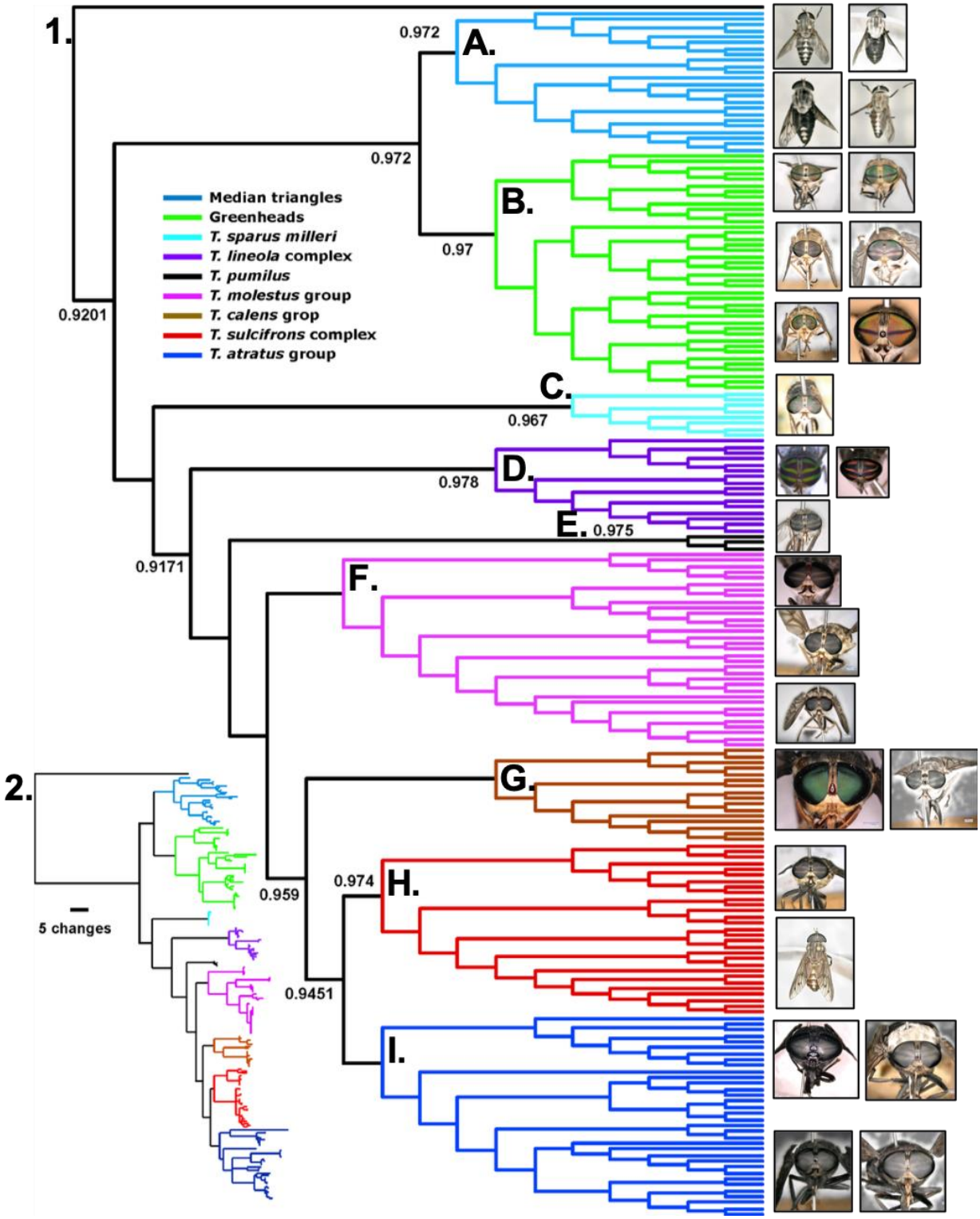

Figure 1 (continued)

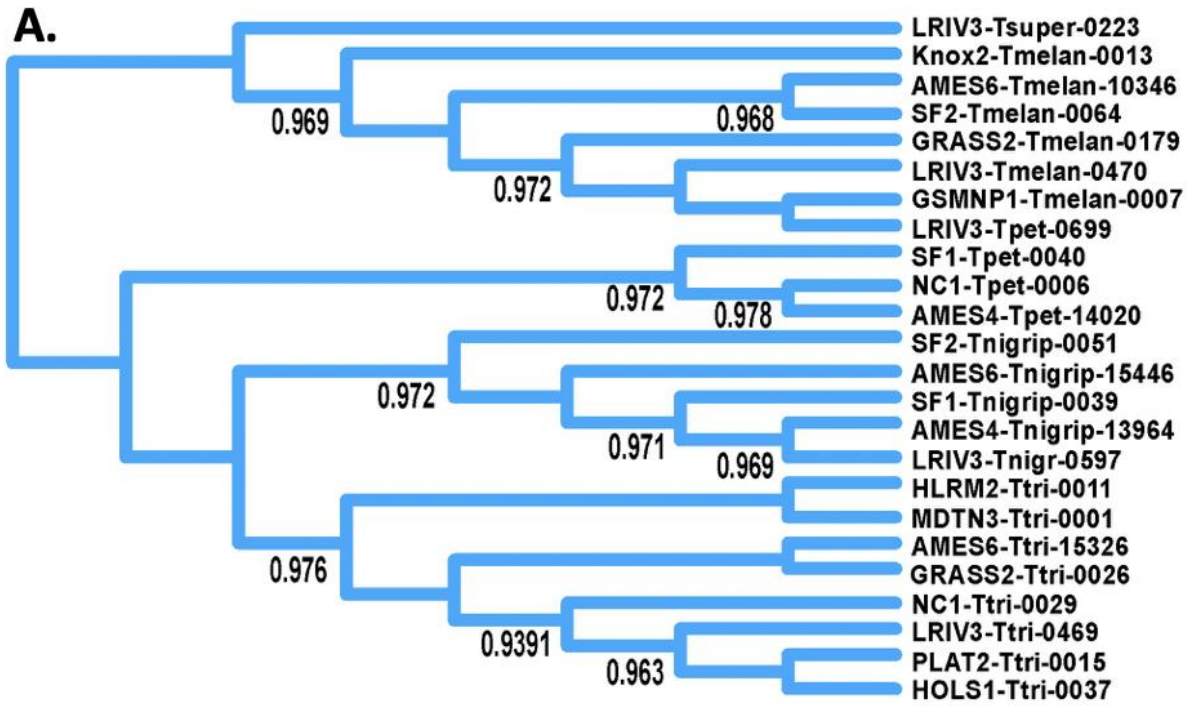

Figure 2. Clade A. Described as the median dorsal pattern tabanids. Species represented are *T. superjumentarius* (-Tsuper-), *T. melanocerus* (-Tmelan-), *T. petiolatus* (-Tpet-), *T. nigripes* (-Tnigr-, -Tnigrip-), and *T. trimaculatus* (-Ttri-). Nodes supported with high posterior probability values ($n \geq 0.95$) are listed.

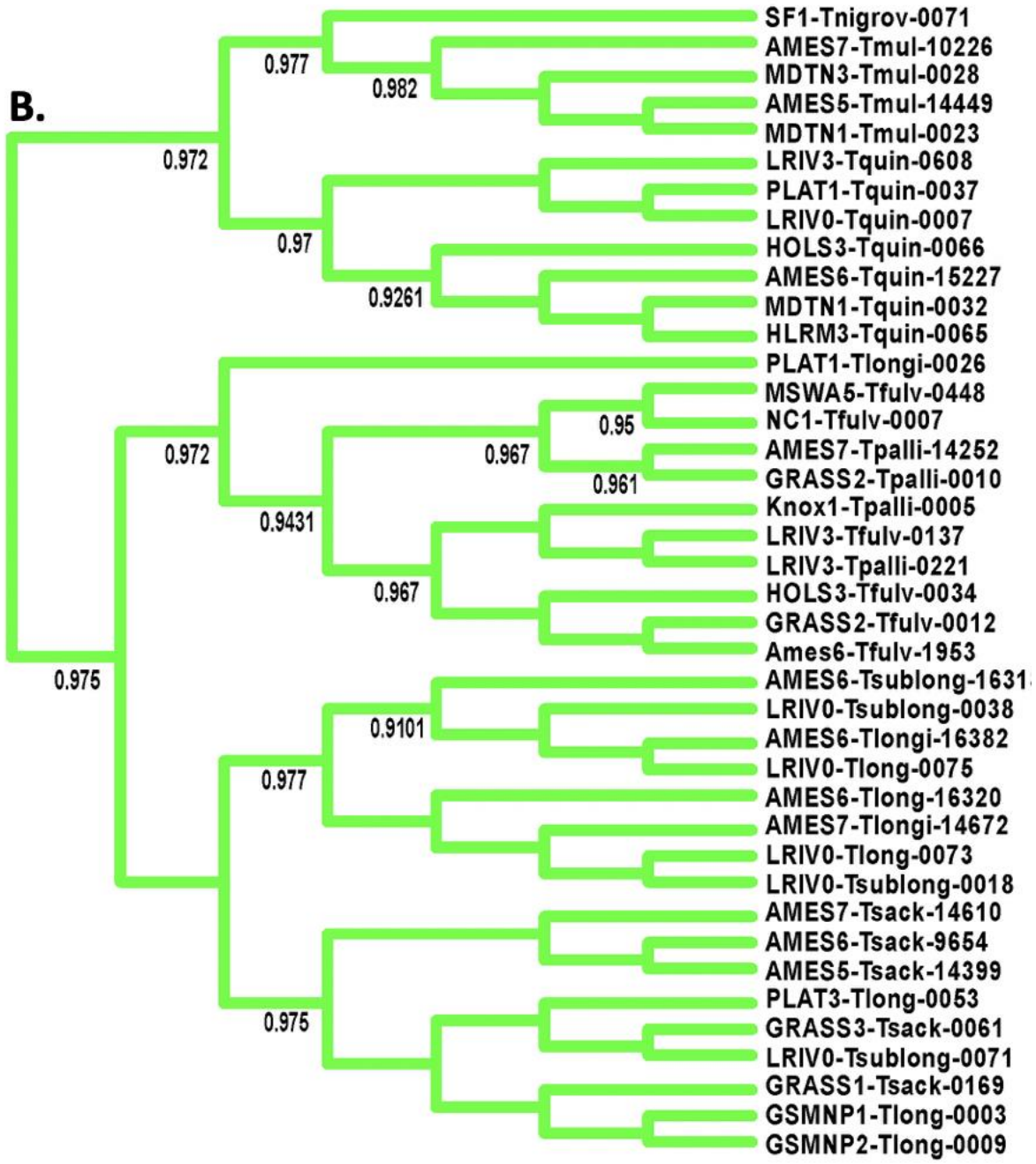

Figure 3. Clade B. Described as greenhead tabanids. Species represented are *T. nigrovittatus* (-Tnigrov-), *T. mularis* (-Tmul-), *T. quinquevittatus* (-Tquin-), *T. longiusculus* (-Tlongi-), *T. fulvulus* (-Tfulv-), *T. pallidescens* (-Tpalli-), *T. sublongus* (-Tsublong-), *T. longus* (-Tlong-), *and T. sackeni* (-Tsack-). Nodes supported with high posterior probability values ($n \geq 0.95$) are listed.

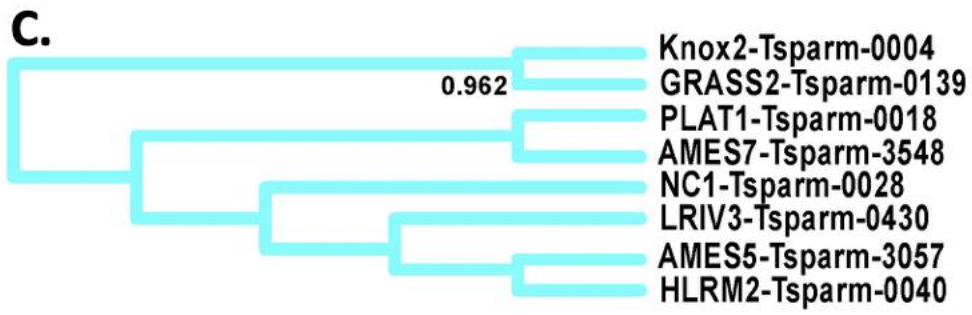

Figure 4. Clade C. Exclusively represents *T. sparus milleri* from 8 specimens and 8 locations. Nodes supported with high posterior probability values ($n \geq 0.95$) are listed.

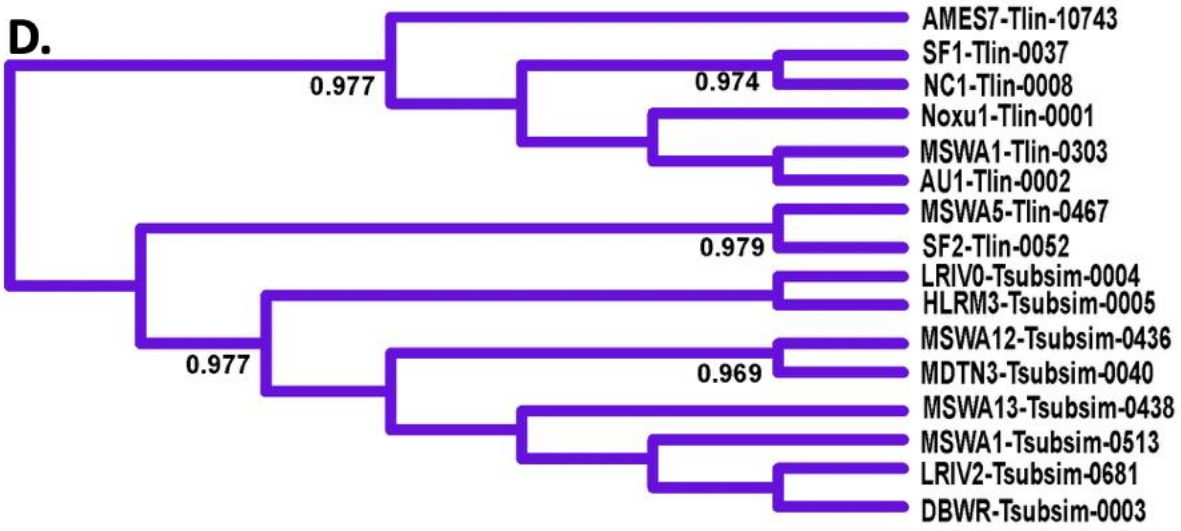

Figure 5. Clade D. Described as the *T. lineola* complex. Species represented are *T. lineola* (-Tlin-) and. *T. subsimilis* (-Tsubsim-). Nodes supported with high posterior probability values ($n \geq 0.95$) are listed.

Figure 6. Clade E. Exclusively represents *T. pumilus* from 3 specimens and 3 locations.

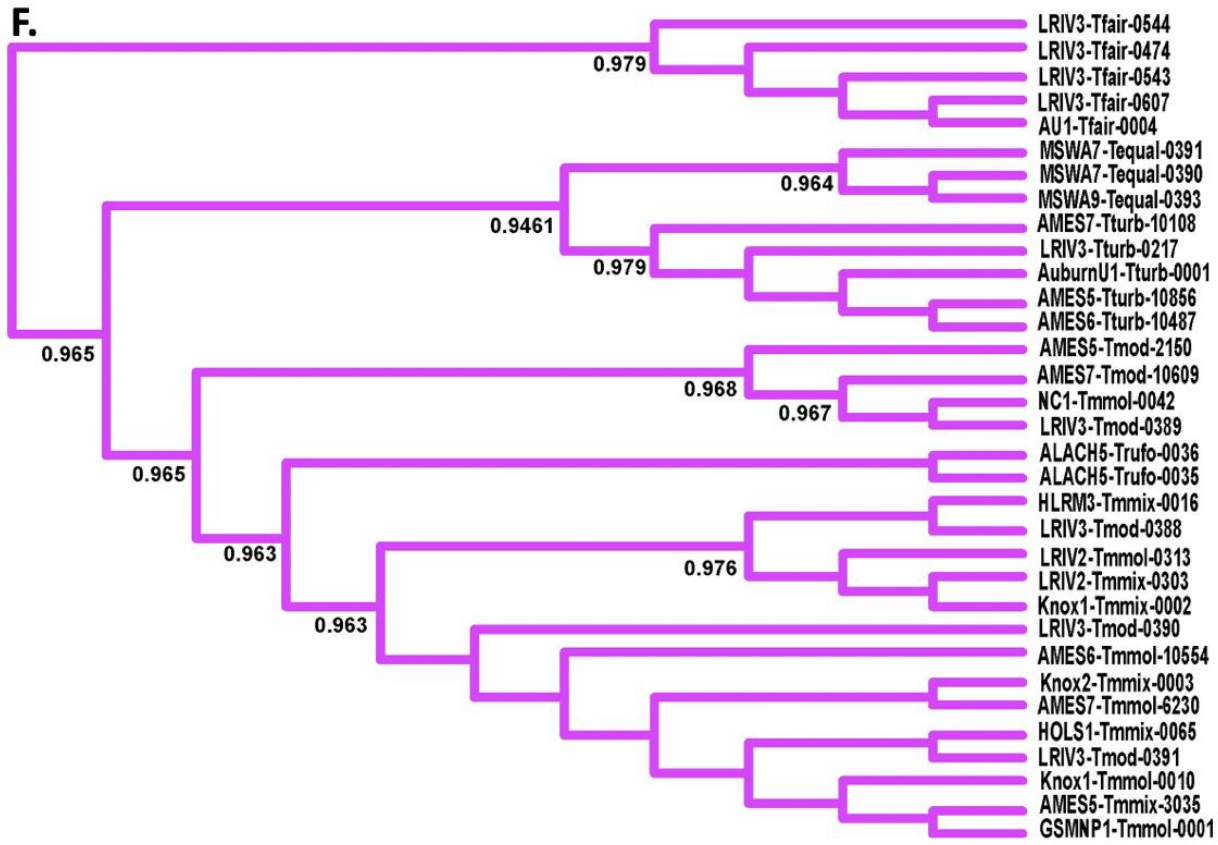

Figure 7. Clade F. Described as the *T. molestus* group. Species represented are *T. Fairchildi* (-Tfair-), *T. equalis* (-Tequal-), *T. turbidus* (-Tturb-), *T. moderator* (-Tmod-), *T. rufofrater* (-Trufo-), *T. mixis* (-Tmmix-), and *T. molestus* (-Tmmol-). Nodes supported with high posterior probability values ($n \geq 0.95$) are listed.

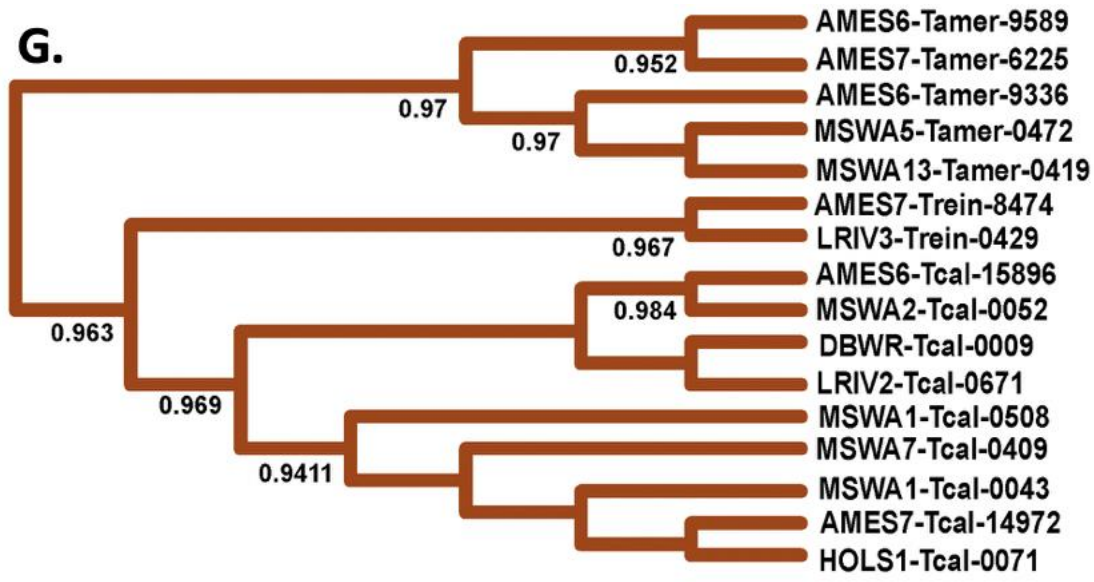

Figure 8. Clade G. Described as the *T. calens* group. Species represented are *T. americanus* (-Tamer-), *T. reinwardtii* (-Trein-), and *T. calens* (-Tcal-). Nodes supported with high posterior probability values (n ≥ 0.95) are listed.

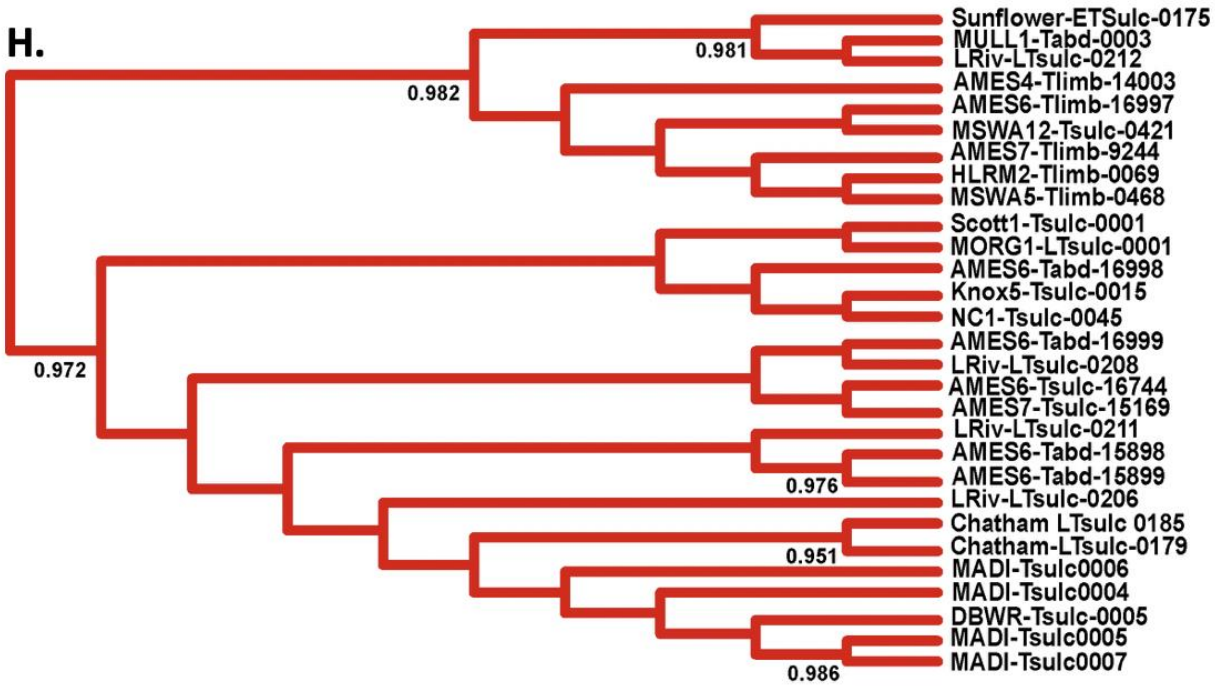

Figure 9. Clade H. Described as the *T. sulcifrons* complex including *T. sulcifrons* (-sulc-), *T. limbatinevris* (-Tlimb-), and *T. abdominalis* (-Tabd-). Nodes supported with high posterior probability values ($n \geq 0.95$) are listed.

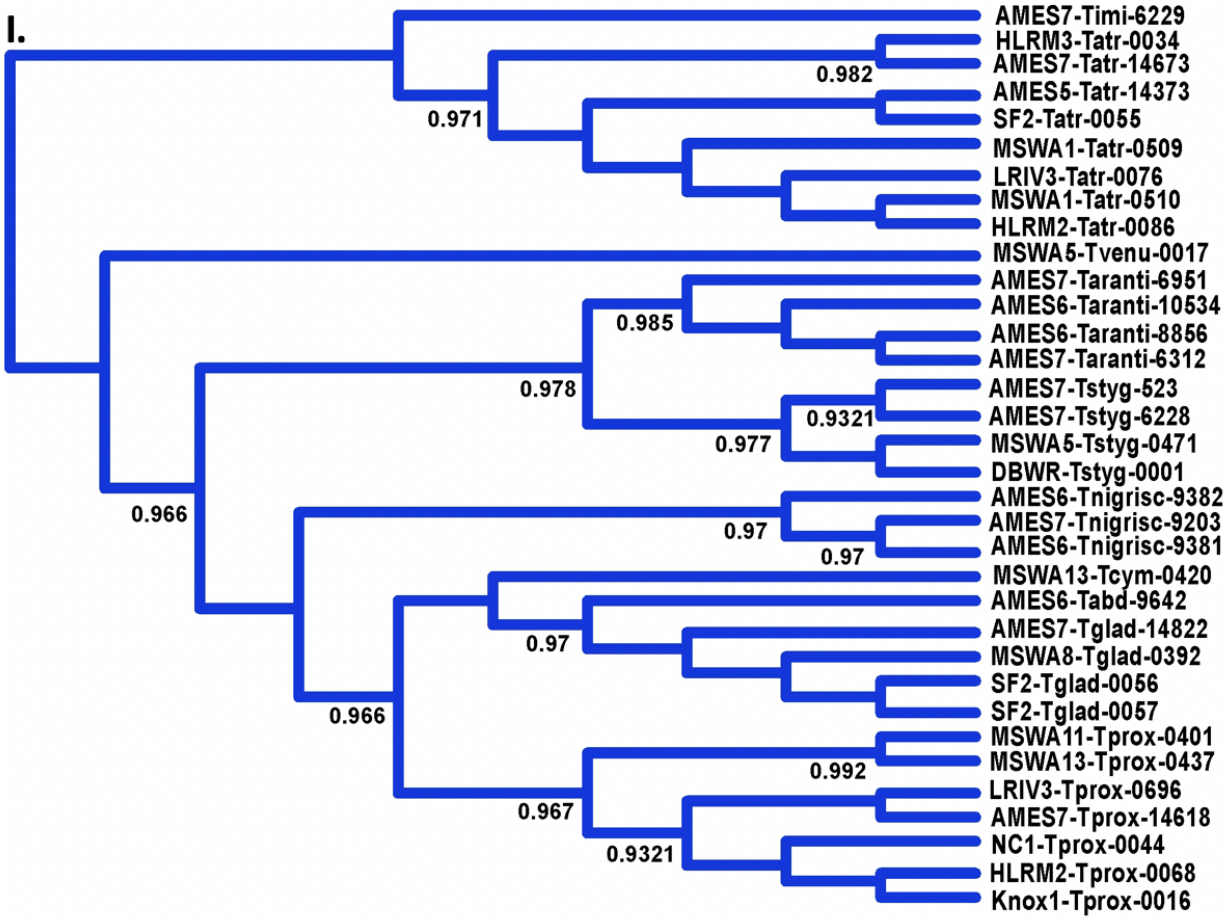

Figure 10. Clade I. Described as the *T. atratus* group. Species represented are *T. imitans* (-Timi-), *T. atratus* (-Tatr-), *T. venustus* (-Tvenu-), *T. aranti* (-Taranti-), *T. stygius* (-Tstyg-), *T. nigriscens* (-Tnigrisc-), *T. cymatomorphous* (-Tcym-), *T. abdominalis* (-Tabd-), *T. gladiator* (-Tglad-), *and T. proximus* (-Tprox-). Nodes supported with high posterior probability values (n ≥ 0.95) are listed.

recent common ancestor of clades F-I. Clade F is described as the *T. molestus group* (Figure 7) and is recovered with low support, but contains species with similar morphologies. Clade G is described as the *T. calens* group (Figure 8) and is strongly supported (0.959) as the sister clade to the most recent common ancestor of clades H and I. Clade H (Figure 9), the sulcifrons complex, is strongly supported (0.9451) as the sister to clade I (Figure 10), the *T. atratus* group.

Conclusions

The phylogeny recovers the lineages of several *Tabanus* species with high support ($n \geq 0.95$) and corroborates morphological similarities among clades A, B, F, G, H and I. Clade A contains horse flies with white triangular markings on the dorsal midline of the female's dark-colored abdomen ranging from tergites 1-6. Clade B is a large group of pestiferous horse flies commonly called greenhead flies for their large green eyes in life. The species of this clade range from a light brown-yellow to orange coloration ranging from 12-14mm in length. Clade F, the *T. molestus* group, contains dark-brownish horse flies with median dorsal patterning generally and wing infuscations. Clade G, the *T. calens* group contains three species with limited morphological similarities, however dried *T. calens* specimens can resemble *T. americanus* specimens when pinned. Clade H, the sulcifrons complex, grouped containing *T. sulcifrons*, *T. abdominalis* and *T. limbatinevris* are closely related with highly resembling morphologies. Clade I, The *T. atratus* was described for predominately large black-colored flies within the group. Most species were recovered and confirmed morphological identifications, however several closely related species were recovered within other species lineages with low to high posterior probability support.

In clade A *T. petiolatus* (LRIV3-Tpet-0699) grouped within the *T. melanocerus* lineage. These species are otherwise distinct with posterior probability support ($n > 0.963$) however, they can be difficult to distinguish as they can be nearly identical in morphology with the exception of a

petiolate first posterior cell and a broadly bridging mid-dorsal triangle from the 2nd tergite to the 1st tergite in *T. petiolatus* (Drees et al. 1980, Goodwin et al. 1985). However, These characters are “not always present” in *T. petiolatus* and “not always absent” in *T. melanocerus* (Teskey 1969).

In clade B, two *T. pallidescens* and *T. fulvulus* lineages were recovered with posterior probability support (n > 0.94). *Tabanus fulvulus* (LRIV3-Tfulv-0137) also grouped within a *T. pallidescens* lineage with low posterior probability support. Explanations for this are *T. pallidescens* and *T. fulvulus* having very similar morphologies. Within the difficult to distinguish greenheads, the longus group (*T. longus, T. sublongus, T. longiusculus, T. sackeni*) had several species group within different species. This is likely due to the dependence on difficult characters to identify specimens when there appears to be a continuum of character states. Among the greenhead flies, COI sequences are most likely to accurately identify *T. mularis* and *T. quinquevittatus* due to monophyly and posterior support (n >0.97).

Tabanus lineola and *T. subsimilis* within the lineola complex in clade D were recovered separately with posterior probability support (n > 0.969). Morphological distinction with dichotomous keys is possible when the dorsal and scutum coloration and basal calli are aligned with the species’ respective holotypes, however there are many individuals that exhibit a continuous character state for these traits by population and location, rendering them not separable consistently. COI data here helped differentiate *T. lineola* and *T. subsimilis* which is important because these species can be difficult to identify and are spatially and temporally sympatric.

In clade F, ancestral nodes were recovered with strong support (n > 0.9461) for the light-reddish brown to dark brown dorsally patterned flies. *Tabanus fairchildi*, *T. equalis*, *T. turbidus* and *T.*

rufofrater are recovered with monophyly and posterior probability support (n > 0.9461). However, *T. molestus* (NC1-Tmmol-0042) grouped within the *T. moderator lineage, T. moderator* (LRIV3-Tmod-0388) *and T. molestus* (LRIV2-Tmmol-0313) *grouped within T. mixis* lineage*,* and *T. moderator* (LRIV3-Tmod-0390) and *T. mixis* (Knox2-Tmmix-0003, HOLS1-Tmmix-0065) were recovered within the *T. molestus* lineage with low posterior probability support. Here the resulting sequence data for *T. molestus* and *T. mixis* indicated that these species are genetically and morphologically similar. Morphologically they are similar with the presence of a distinct brown and a white pilosity observed in *T. mixis* and *T. molestus* respectively, however *T. molestus* is also extremely similar morphologically to *T. moderator*. *Tabanus molestus*, *T. mixis*, and *T. moderator* are considered as distinct species despite their morphological similarities (Nalen et al. 2015). It is likely that more molecular data is necessary to distinguish these closely related species.

Clade H contains notorious livestock pests within the *sulcifrons complex*, including *T. abdominalis* and *T. limbatinevris*, with two *T. sulcifrons* lineages recovered with support (0.9451) and specimens recovered within different lineages (Drees et al. 1980, Goodwin et al. 1985). *Tabanus sulcifrons* (MSWA12-Tsulc-0421) was recovered with the *T. limbatinevris* lineage and *T. abdominalis* (MULL1-Tabd-0003, AMES6-Tabd-16998, AMES6-Tabd-16999) were recovered within *T. sulcifrons* with low support. *Tabanus sulcifrons* exhibits considerable morphological variation and will require more molecular data to resolve species and lineage questions within the complex. In the sister clade I, the majority of this group of large fly species were recovered in support of morphological data and posterior probability. However, *T. abdominalis* was recovered within the *T. gladiator* lineage with support (0.97). An explanation for this is *Tabanus gladiator* is considered another member of the *sulcifrons complex* although it is more distinctive on average morphologically due to the lavender-tinted thorax among other features (Goodwin et

al. 1985). Overall, members within the *sulcifrons complex* exhibit considerable morphological variation and require more molecular data to resolve taxonomic and lineage questions.

In the phylogeny, low probability scores (n < 0.90) are likely a result of a combination of excluded sequences of species that are part of the *Tabanus* evolutionary history but were not collected, the use of one mitochondrial marker, and the need for more gene sequences to examine the evolutionary and gene history of *Tabanus* more extensively. Another possibility for disparities and error are misidentifications of species that are very closely related and difficult to distinguish with morphology. Measures to limit this have taken place as identifications were all verified and confirmed. With species distinction as a priority, clades with lowest support contained species with relatively stable adult morphology that are distinguishable with well-constructed keys and many species were still able to be identified with COI.

At the moment, data produced in this study can help identify 40 *Tabanus* species, many of which are major pests in the Southeastern U.S., as well as contribute to the barcode of life initiative and knowledge of insect diversity, species diversification, and evolution (Costa and Carvalho 2010). Molecular identification to incriminate vectors and survey pest populations is increasingly a practical and viable option in the absence of taxonomists and other forms of direct identification support. Molecular identification of immature and adult *Tabanus* stages represented is achievable with the data provided in this manuscript and will allow for characterization of larval habitat and development with surveillance efforts. Learning more about the life cycles of horse flies offers an opportunity to exploit their behavior to sustainably and adequately control pest populations.

Chapter 2

Improving the identification of *Tabanus* flies (Diptera: Tabanidae) with environmental niche modeling

Abstract

Lack of identification tools for *Tabanus* have allowed this important group of biting flies to be understudied and remain a persistent economic problem to livestock producers and a health and welfare problem to livestock. Horse fly ecology is well described from specimen collections but lacks support from analytical data. Our objective was to produce environmental niche models (ENM) that are useful as identification tools for horse flies presuming that each niche model potentially characterizes a species by its distribution and niche. Models were produced for the most pervasive horse flies collected in this study in the Southeastern U.S. (*T. fulvulus* Wiedemann*, T. lineola* Fabricius*, T. subsimilis* Bellardi*, T. quinquevittatus* Wiedemann*, T. sparus milleri* Whitney, and *T. sulcifrons* Macquart). Maximum entropy models were produced with Maxent 3.4.1 with county centroid record data and with GPS-data coordinates obtained from trapping and collection sites. Climate data were obtained from PRISM climate group (1980-2010) and remotely sensed data from the USGS (2000-2016). Distributions for the horse fly species modeled from GPS-data range across the state of Tennessee and 17 surrounding states in the Eastern U.S. High relative humidity and temperature range (continentality) were analyzed as the most important environmental variables to model contribution and the GPS-data ENMs aligned within the broad historical collection range as shown with the county centroid ENMs. This research provides distinguishing distributional data for horse flies and facilitates further horse fly investigations of feeding behavior, host specificity, seasonality, pathogen transmission, and life history.

Introduction

Repeated bites and blood meals by horse flies (Diptera: Tabanidae) cause blood loss, behavioral disturbance, and stress to livestock that results in animals losing weight gain and operations losing profit (Perich et al. 1986, Foil and Hogsette 1994). Throughout the summer adult horse fly populations peak, and they can also be alarming pests for people during outdoor

activities and recreation (Teskey and Pechuman 1983). Control of horse flies is currently ineffective because little is known about the ecology of the different involved *Tabanus* species as it relates to their life cycle. Distributions of eastern U.S. *Tabanus* flies previously mapped are not descriptive enough to target specific environments for study or collection and provide limited ecological insight (Teskey and Pechuman 1983). Distribution and niche data provide background information for species description and can be especially useful with diverse and closely related organisms as a means of identification (Raxworthy et al. 2007). For these reasons, horse flies are ideal candidates for environmental niche modeling (ENM) which can be used to better understand their ecology as pests and vectors, life history, and to improve species identification.

Immature and oviposition habitats of *Tabanus* are speculated to be among the most important criteria for immature and adult occurrence and include the margins of bodies of water with emergent vegetation, driftwood, drained soil or periodically flooded areas with aquatic and semiaquatic components (Jones and Anthony 1964, Goodwin et al. 1985). This wide range of habitats may be suitable to describe the distributions of some suspected generalist horse flies (*T. subsimilis* Bellardi, *T. lineola* Fabricius, *T. atratus* Fabricius), but these habitats do not precisely describe the niches and distributions of the many species that are likely very diverse and overlapping. To properly identify *Tabanus* habitat and to improve identification methods for *Tabanus* research our objective was to use climatic and land cover variables with *Tabanus* occurrence data and fit them into a maximum entropy environmental niche model.

A study in Ecuador analyzed the distributions of three Tabanidae species with maximum entropy modeling. Each species had a unique suitability range that matched with distinctive abiotic factors such as altitude (Cárdenas et al. 2009). The same approach Is useful in ascertaining ecological data of Nearctic Tabanidae. Maxent maximum entropy environmental

niche modeling uses GPS presence data to produce the environmental suitability of a species as the maximum entropy probability distribution displayed as pixel values in the projected study range (Raxworthy et al. 2007). Geographic coordinates should be recorded from the precise location of collection to be used in modeling. In the case of county records with no coordinates recorded, a county centroid approach can be used in order to obtain useable occurrence data but should not be interpreted the as precise coordinates. One-kilometer resolution suitability ranges improve ecological knowledge of identified species with respect to their physical environment and can characterize parapatric species geographic distinctions more precisely (Raxworthy et al. 2007). Maxent is the ideal modeling method for the needs of this study because it does not require absence data and limits commission error (Ferrier et al. 2006). Environmental suitability is calculated as a function of how known species occurrence records correspond with the averages of the environmental and land cover features in the form of raster values within the training area of the model (Phillips et al. 2006). Spatially unique horse fly species modeled with abiotic predictor variables will estimate the distribution and niche suitability. Producers, veterinarians, and researchers can then use these models to help identify the pestiferous livestock problem.

Materials and Methods

Tabanus collections and occurrence data

Methods for material and specimen collection were previously described in Chapter 1. Briefly, occurrence data came from a combination of sources which included active *Tabanus* collections and museum identifications. Specimens were physically collected with the H trap (Bite-Lite, Bethel, CT), the NZI trap (Rincon-Vitova, Ventura, CA), the I trap (details described in Davis et al. unpublished), and through incidental occurrences. Traps were operated in Alabama (1 site), Arkansas (1 site), Florida (14 sites), Mississippi (14 sites), North Carolina (2 sites), and Tennessee (42 sites) at least 1 kilometer apart when possible (Table 2). The GPS coordinates

of these collections were used to create ENMs for each species. Additionally, museum county records were obtained (Table 3) from the University of Tennessee Entomology and Plant Pathology insect museum, the Cornell Insect Collection (CUIC), the University of AUMNH Entomology collection, and the Florida State Collection of Arthropods (FSCA) in Gainesville, FL. County centroid geographic coordinates were obtained with ArcMap tools and were used to generate county-level ENMs.

Environmental niche modeling was performed for species with sufficient collection records, which I define as at least 20 sites more than 1km apart (Pearson et al. 2007). Two ENM models were generated, the first with GPS trapping and collection records data and the second with county-centroid data. Six *Tabanus* species (*T. fulvulus* Wiedemann*, T. lineola* Fabricius*, T. subsimilis* Bellardi*, T. quinquevittatus* Wiedemann*, T. sparus milleri* Whitney, and *T. sulcifrons* Macquart) were identified with sufficient records and were also identified as the most pervasive fly species of the study area which was indicated by their widespread distribution and abundance.

Environmental variables

Climatic and land cover predictor variables were selected based on their contribution to model variation in previous studies and with respect to biological relevance (Cárdenas et al. 2009). Environmental predictor variables were pre-selected based on each individual variable's performance in the distributions of neotropical tabanids and previously reported biological importance to *Tabanus* species in Ecuador (Cárdenas et al. 2009). Eight climatic variables collected from 1980 to 2010 (mean annual temperature, continentality, mean annual relative humidity, mean annual precipitation, mean summer precipitation, summer heat moisture index, degree days above 18ºC, number of frost-free days) were downloaded using the PRISM Climate North America dataset (C. Daly et al. 2000). The mean values of four land

Table 2. Species occurrence data of pervasive horse flies. Species occurrence data of pervasive horse flies for GPS-data models. Trap type or collection method is listed for each species occurrence point and the resulting geographic coordinates.

Species	Trap/Method	Location	Latitude	Longitude
T. fulvulus	NZI	Little River Farm Blount co., TN	35.76356	-83.83694
	H trap	Knoxville, TN	35.90832	-83.86278
	Net	Ames Plantation Fayette, TN	35.1134	-89.20611
	H trap	Little River Farm Blount co., TN	35.77021	-83.85472
	H trap	Little River Farm Blount co., TN	35.761	-83.83944
	NZI	Little River Farm Blount co., TN	35.76942	-83.84417
	NZI	Holston Farm Knoxville, TN	35.96203	-83.85861
	I trap	Holston Farm Knoxville, TN	35.96078	-83.85444
	I trap	Ames Plantation Ames7	35.13612	-89.28083
	H trap	Ames Plantation Fayette co., TN	35.13442	-89.22278
	NZI	Ames Plantation Fayette co., TN	35.14392	-89.24444

Table 2 (continued)

Species	Trap/Method	Location	Latitude	Longitude
T. fulvulus	NZI	Plateau R.E.C. Cumberland, TN	36.02039	-85.1275
	H trap	Plateau R.E.C. Cumberland, TN	36.00996	-85.11806
	I trap	Grasslands Farms Cumberland, TN	35.84824	-85.06194
	H trap	Grasslands Farms Cumberland, TN	35.83811	-85.06917
	I trap	Highland Rim R.E.C. Springfield, TN	36.47844	-86.83889
	Malaise	Sunflower, Mississippi	33.43086	-90.69535
	Malaise	Bienville, Mississippi	32.31722	-89.48983
	H trap	Sandford, NC	35.61199	-79.16476
	H trap	Holston Farm Knoxville, TN	35.95742	-83.86
	Horsepal	Ames Plantation Fayette co., TN	35.11252	-89.21611
	H trap	Ames Plantation Fayette co., TN	35.11056	-89.2125
	H trap	Ames Plantation Fayette co., TN	35.10889	-89.21611

Table 2 (continued)

Species	Trap/Method	Location	Latitude	Longitude
T. fulvulus	H trap	Ames Plantation Fayette co., TN	35.11395	-89.21722
	Net	Ames Plantation Fayette co., TN	35.11244	-89.2025
T. lineola	H trap	CMAVE USDA Alachua co., FL	29.635719	-82.360055
	NZI	Austin Cary Forest Alachua co., FL	29.599456	-82.349501
	NZI	Austin Cary Forest Alachua co., FL	29.74158	-82.216922
	I trap	Little River Farm Blount co., TN	35.7702119	-83.854722
	H trap	Little River Farm Blount co., TN	35.7610042	-83.839444
	I trap	Plateau R.E.C. Cumberland co., TN	36.0191961	-85.135556
	I trap	Grasslands Farms Cumberland co., TN	35.8381078	-85.069167
	NZI	Grasslands Farms Cumberland co., TN	35.8482422	-85.061944
	I trap	Ames Plantation Fayette Co., TN	35.1361197	-89.280833

Table 2 (continued)

Species	Trap/Method	Location	Latitude	Longitude
T. lineola	NZI	Ames Plantation Fayette Co., TN	35.1439242	-89.244444
	H trap	Ames Plantation Fayette Co., TN	35.1344228	-89.222778
	malaise	Bolivar co., MS	33.704805	-89.644028
	malaise	Sunflower co., MS	33.430859	-90.695346
	Malaise	Tallahatchie co., MS	33.757902	-90.149618
	NZI	Auburn, Alabama	32.5794469	-85.501389
	H trap	Panama City Beach, FL	30.255695	-85.897195
	H trap	Panama City Beach, FL	30.217642	-85.851888
	H trap	Sandford, NC	35.611989	-79.164758
	Malaise	Dale Bumpers White River Wildlife Refuge, Arkansas	34.35647	-91.120933
	Net	Noxubee Wildlife Refuge, MS	33.2708333	-88.782222
	Horsepal	Ames Plantation Fayette Co., TN	35.1125203	-89.216111
	H trap	Ames Plantation Fayette Co., TN	35.1139525	-89.217222
	H trap	Ames Plantation Fayette Co., TN	35.1105603	-89.2125
	H trap	Ames Plantation Fayette Co., TN	35.1088864	-89.216111

Table 2 (continued)

Species	Trap/Method	Location	Latitude	Longitude
T. subsimilis	NZI	Little River Farm Blount co., TN	35.7635589	-83.836944
	NZI	Little River Farm Blount co., TN	35.7694169	-83.844167
	I trap	Little River Farm Blount co., TN	35.7610042	-83.839444
	NZI	Little River Farm Blount co., TN	35.7702119	-83.854722
	Malaise	Sunflwer co., MS	33.430859	-90.695346
	Malaise	Sunflwer co., MS	33.462079	-90.707222
	I trap	Holston Farm Knoxville, TN	35.9607758	-83.854444
	H trap	Plateau R.E.C. Cumberland co., TN	36.009955	-85.118056
	I trap	Grasslands Farms Cumberland co., TN	35.8381078	-85.069167
	NZI	Grasslands Farms Cumberland co., TN	35.8401728	-85.053333
	NZI	Highland Rim R.E.C. Springfield, TN	36.4672503	-86.838056
	I trap	Highland Rim R.E.C. Springfield, TN	36.4784394	-86.838889
	H trap	Highland Rim R.E.C. Springfield, TN	36.4721531	-86.818333

Table 2 (continued)

Species	Trap/Method	Location	Latitude	Longitude
T. subsimilis	I trap	Ames Plantation Fayette Co., TN	35.1361197	-89.280833
	NZI	Ames Plantation Fayette Co., TN	35.1439242	-89.244444
	H trap	Ames Plantation Fayette Co., TN	35.1344228	-89.222778
	NZI	Holston Farm Knoxville, TN	35.9620286	-83.858611
	Malaise	Bolivar co., MS	33.704805	-89.644028
	Malaise	Montgomery co., MS	33.439167	-89.644028
	Malaise	James, MS	33.160361	-91.054065
	malaise	Rolling Fork, MS	32.91696	-90.92081
	malaise	Tallahatchie co., MS	33.760588	-90.150779
	malaise	Tallahatchie co., MS	33.757902	-90.149618
	H trap	CMAVE USDA Alachua co., FL	29.635719	-82.360055
	NZI	Austin Cary Forest Alachua co., FL	29.74158	-82.216922
	Net	Alachua co., FL	29.6805697	-82.311389
	H trap	Holston Farm Knoxville, TN	35.957415	-83.86
	Malaise	Dale Bumpers White River Wildlife Refuge, Arkansas	34.35647	-91.120933

Table 2 (continued)

Species	Trap/Method	Location	Latitude	Longitude
T. subsimilis	H trap	Grasslands Farms Cumberland co., TN	35.8482422	-85.061944
	I trap	Middle TN R.E.C. Spring Hill, TN	35.71691	-86.951944
	Horsepal	Ames Plantation Fayette Co., TN	35.1139525	-89.217222
	H trap	Ames Plantation Fayette Co., TN	35.1125203	-89.216111
	H trap	Ames Plantation Fayette Co., TN	35.1105603	-89.2125
	H trap	Ames Plantation Fayette Co., TN	35.1088864	-89.216111
T. quinquevittatus	NZI	Little River Farm Blount co., TN	35.7635589	-83.836944
	NZI	Little River Farm Blount co., TN	35.7610042	-83.839444
	H trap	Little River Farm Blount co., TN	35.7694169	-83.844167
	H trap	Little River Farm Blount co., TN	35.7702119	-83.854722
	NZI	Holston Farm Knoxville, TN	35.9620286	-83.858611
	I trap	Holston Farm Knoxville, TN	35.9607758	-83.854444

Table 2 (continued)

Species	Trap/Method	Location	Latitude	Longitude
T. quinquevittatus	H trap	Holston Farm Knoxville, TN	35.957415	-83.86
	NZI	Plateau R.E.C. Cumberland co., TN	36.0203939	-85.1275
	I trap	Plateau R.E.C. Cumberland co., TN	36.0191961	-85.135556
	H trap	Plateau R.E.C. Cumberland co., TN	36.009955	-85.118056
	NZI	Grasslands Farms Cumberland co., TN	35.8401728	-85.053333
	H trap	Grasslands Farms Cumberland co., TN	35.8482422	-85.061944
	I trap	Grasslands Farms Cumberland co., TN	35.8381078	-85.069167
	I trap	Middle TN R.E.C. Spring Hill, TN	35.7145947	-86.970278
	I trap	Middle TN R.E.C. Spring Hill, TN	35.7145947	-86.970278
	NZI	Highland Rim R.E.C. Springfield, TN	36.4672503	-86.838056
	I trap	Highland Rim R.E.C. Springfield, TN	36.4784394	-86.838889

Table 2 (continued)

Species	Trap/Method	Location	Latitude	Longitude
T. quinquevittatus	H trap	Highland Rim R.E.C. Springfield, TN	36.4721531	-86.818333
	Net	Ames Plantation Fayette Co., TN	35.1361197	-89.280833
	H trap	Ames Plantation Fayette Co., TN	35.1344228	-89.222778
	NZI	Ames Plantation Fayette Co., TN	35.1439242	-89.244444
	H trap	Ames Plantation Fayette Co., TN	35.1139525	-89.217222
	Horsepal	Ames Plantation Fayette Co., TN	35.1125203	-89.216111
	H trap	Ames Plantation Fayette Co., TN	35.1105603	-89.2125
	H trap	Ames Plantation Fayette Co., TN	35.1088864	-89.216111
T. sparus milleri	NZI	Little River Farm Blount co., TN	35.7635589	-83.836944
	I trap	Little River Farm Blount co., TN	35.7694169	-83.844167
	H trap	Little River Farm Blount co., TN	35.7610042	-83.839444

Table 2 (continued)

Species	Trap/Method	Location	Latitude	Longitude
T. sparus milleri	H trap	Little River Farm Blount co., TN	35.7702119	-83.854722
	Net	Knoxville, TN	36.08718	-83.8201
	Net	Knoxville, TN	35.99298	-83.86488
	NZI	Ames Plantation Fayette Co., TN	35.1439242	-89.244444
	H trap	Ames Plantation Fayette Co., TN	35.1344228	-89.222778
	I trap	Ames Plantation Fayette Co., TN	35.1361197	-89.280833
	I trap	Plateau R.E.C. Cumberland co., TN	36.0191961	-85.135556
	H trap	Plateau R.E.C. Cumberland co., TN	36.009955	-85.118056
	NZI	Grasslands Farms Cumberland co., TN	35.8401728	-85.053333
	H trap	Grasslands Farms Cumberland co., TN	35.8482422	-85.061944
	I trap	Grasslands Farms Cumberland co., TN	35.8381078	-85.069167
	I trap	Highland Rim R.E.C. Springfield, TN	36.4784394	-86.838889
	H trap	Sandford, NC	35.611989	-79.164758

Table 2 (continued)

Species	Trap/Method	Location	Latitude	Longitude
T. sulcifrons	Horsepal	Ames Plantation Fayette Co., TN	35.1125203	-89.216111
	H trap	Ames Plantation Fayette Co., TN	35.1139525	-89.217222
	H trap	Ames Plantation Fayette Co., TN	35.1105603	-89.2125
	H trap	Ames Plantation Fayette Co., TN	35.1088864	-89.216111
	NZI	Little River Farm Blount co., TN	35.7610042	-83.839444
	H trap	Little River Farm Blount co., TN	35.7702119	-83.854722
	H trap	Knoxville, TN	35.9083194	-83.862778
	Net	Chapman Highway, Knoxville, TN	35.906375	-83.836944
	Net	Lauderdale Co., TN	35.701025	-89.656111
	Net	Frozen Head Park, Morgan Co., TN	36.127472	-84.501125
	H trap	Sandford, NC	35.611989	-79.164758
	H trap	Holston Farm Knoxville, TN	35.957415	-83.86
	I trap	Grasslands Farms Cumberland co., TN	35.8381078	-85.069167

Table 2 (continued)

Species	Trap/Method	Location	Latitude	Longitude
T. sulcifrons	H trap	Grasslands Farms Cumberland co., TN	35.8482422	-85.061944
	net	Middle TN R.E.C. Spring Hill, TN	35.7175603	-86.958056
	I trap	Middle TN R.E.C. Spring Hill, TN	35.71691	-86.951944
	I trap	Highland Rim R.E.C. Springfield, TN	36.4784394	-86.838889
	Net	Scott Co., TN	36.3537139	-84.585556
	Net	Hef USA Chattanooga, TN	35.039415	-85.322778
	H trap	Ames Plantation Fayette Co., TN	35.1088864	-89.216111
	H trap	Ames Plantation Fayette Co., TN	35.1139525	-89.217222
	H trap	Ames Plantation Fayette Co., TN	35.1105603	-89.2125
	NZI	Ames Plantation Fayette Co., TN	35.1439242	-89.244444
	H trap	Ames Plantation Fayette Co., TN	35.1344228	-89.222778
	Net	Ames Plantation Fayette Co., TN	35.1361197	-89.280833

Table 2 (continued)

Species	Trap/Method	Location	Latitude	Longitude
T. sulcifrons	Horsepal	Ames Plantation Fayette Co., TN	35.1125203	-89.216111
	I trap	ETREC Blount Unit TN	35.8418364	-83.949167

Table 3. County occurrence records of pervasive horse flies. Occurrence of pervasive horse flies by county, state records. The latitude and longitude listed correspond to the centroid geographic coordinates of each county.

Species	County	State	Latitude	Longitude
T. fulvulus	Coffee	TN	35.488759	-86.078219
	Cumberland	TN	35.952398	-84.994761
	Monroe	TN	35.447666	-84.249786
	Fayette	TN	35.196993	-89.413803
	Stewart	TN	36.511756	-87.851548
	Madison	TN	35.606056	-88.833424
	Blount	TN	35.688185	-83.922973
	Knox	TN	35.992727	-83.937721
	Robertson	TN	36.52753	-86.869377
	Covington	AL	31.243987	-86.448721
	Charleston	SC	32.800458	-79.94248
	Georgetown	SC	33.41753	-79.300812
	Oktibbeha	MS	33.422313	-88.876151
	Tuscaloosa	AL	33.290202	-87.52286
	Lancaster	PA	40.041992	-76.250198
	Burlington	NJ	39.875786	-74.663006
	Houston	AL	31.158193	-85.296398
	Conecuh	NY	31.428293	-86.992029
T. lineola	Cumberland	TN	35.952398	-84.994761
	Madison	TN	35.606056	-88.833424
	Fayette	TN	35.196993	-89.413803

Table 3 (continued)

Species	County	State	Latitude	Longitude
T. lineola	Monroe	TN	35.447666	-84.249786
	Knox	TN	35.992727	-83.937721
	Blount	TN	35.688185	-83.922973
	East Baton Rouge Parish	LA	30.544002	-91.093174
	Carteret	NC	34.858313	-76.526967
	Robertson	TN	36.52753	-86.869377
	Maury	TN	35.615696	-87.077763
	Coffee	TN	35.488759	-86.078219
	Lake	IL	42.326444	-87.436118
	Columbus	NC	34.260471	-78.636378
	Johnston	NC	35.513405	-78.367267
	Cohier	FL	26.118713	-81.400884
	Bladen	NC	34.591949	-78.539513
	Beaufort	NC	35.482313	-76.842014
	Oktibbeha	MS	33.422313	-88.876151
	Burlington	NJ	39.875786	-74.663006
	Lancaster	SC	40.041992	-76.250198
	Union	IL	37.475104	-89.252875
	Highlands	FL	27.342627	-81.340921
	Collier	FL	26.118713	-81.400884
	Highlands	FL	27.342627	-81.340921
	Dade	FL	25.610494	-80.499045

Table 3 (continued)

Species	County	State	Latitude	Longitude
T. lineola	Collier	FL	26.118713	-81.400884
	Richland	SC	38.71155	-88.085698
	Madison	AL	34.764238	-86.55108
	Baldwin	AL	30.659218	-87.746067
	Conecuh	AL	31.428293	-86.992029
T. subsimilis	Madison	TN	35.606056	-88.833424
	Fayette	TN	35.196993	-89.413803
	Monroe	TN	35.447666	-84.249786
	Knox	TN	35.992727	-83.937721
	Blount	TN	35.688185	-83.922973
	Wayne	NC	35.362741	-78.004826
	Johnson	IL	37.460815	-88.882962
	Gates	NC	36.442135	-76.702355
	Maury	TN	35.615696	-87.077763
	Greene	TN	36.178998	-82.847746
	Lake	TN	36.333905	-89.485537
	Robertson	TN	36.52753	-86.869377
	Hardeman	TN	35.218131	-88.989037
	Osage	OK	36.62468	-96.408385
	Rockingham	NH	42.98936	-71.099437
	Duval	FL	30.335245	-81.648113
	Fairfax	VA	38.833743	-77.276117
	Willacy	TX	26.481092	-97.584224

Table 3 (continued)

Species	County	State	Latitude	Longitude
T. subsimilis	Lee	AL	32.604064	-85.353048
T. quinquevittatus	Cumberland	TN	35.952398	-84.994761
	Monroe	TN	35.447666	-84.249786
	Knox	TN	35.992727	-83.937721
	Fayette	TN	35.196993	-89.413803
	Madison	TN	35.606056	-88.833424
	Blount	TN	35.688185	-83.922973
	Carteret	NC	34.858313	-76.526967
	Wake	NC	35.789846	-78.650624
	Greene	TN	36.178998	-82.847746
	Robertson	TN	36.52753	-86.869377
	Adams	PA	39.869471	-77.21773
	Kent	MD	39.239177	-76.1242
	Oneida	NY	43.242727	-75.434282
	Bucks	PA	40.336887	-75.10706
	Lake	IL	42.326444	-87.436118
	Callaway	MO	38.835966	-91.924089
	Tippecanoe	IN	40.38926	-86.893943
	Adams	PA	39.869471	-77.21773
	Lancaster	PA	40.041992	-76.250198
	Morgan	OH	39.624946	-81.861699
T. sparus milleri	Monroe	TN	35.447666	-84.249786
	Madison	TN	35.606056	-88.833424

Table 3 (continued)

Species	County	State	Latitude	Longitude
T. sparus milleri	Blount	TN	35.688185	-83.922973
	Monroe	TN	35.447666	-84.249786
	Lancaster	SC	40.041992	-76.250198
	Johnston	NC	35.513405	-78.367267
	Oktibbeha	MS	33.422313	-88.876151
	Burlington	NJ	39.875786	-74.663006
	Champaign	OH	40.132759	-83.767543
	Franklin	OH	39.969447	-83.008258
	Lawrence	OH	38.603866	-82.517186
	Pinellas	FL	27.903122	-82.739518
	Cumberland	TN	35.952398	-84.994761
	Fayette	TN	35.196993	-89.413803
	Coffee	TN	35.488759	-86.078219
	Chester	TN	35.416639	-88.605505
	Greene	TN	36.178998	-82.847746
	Knox	TN	35.992727	-83.937721
T. sulcifrons	DeKalb	AL	34.4597996	-85.804109
	Wilcox	AL	31.9893044	-87.308195
	Duval	FL	30.3315733	-81.670843
	Putnam	FL	29.6086505	-81.74431
	Bryan	GA	32.0144705	-81.443638
	Talbot	GA	32.6995	-84.533009
	Natchitoches	LA	31.7235371	-93.096224

Table 3 (continued)

Species	County	State	Latitude	Longitude
T. sulcifrons	Ouachita	LA	32.4783198	-92.154865
	Washington	MS	33.2837811	-90.947487
	Bladen	NC	34.6145875	-78.563639
	Columbus	NC	34.2655819	-78.655021
	Hyde	NC	35.5304858	-76.250805
	Pasquotank	NC	36.2954736	-76.283987
	Sampson	NC	34.9915502	-78.371388
	Pickens	SC	34.8874779	-82.725309
	Coffee	TN	35.4906187	-86.074753
	Monroe	TN	35.4426465	-84.252734
	Morgan	TN	36.1350081	-84.649198
	Madison	GA	34.1277848	-83.209036
	Oglethorpe	GA	33.8806677	-83.080712
	Putnam	GA	33.3217667	-83.372794
	Lancaster	VA	37.7345211	-76.46322
	Elmore	AL	32.5966467	-86.149159
	Houston	AL	31.1531998	-85.302472
	Jackson	AL	34.7794524	-85.999355
	Madison	AL	34.7630899	-86.550226
	Arkansas	AR	34.290809	-91.374911
	Tangipahoa	LA	30.6266308	-90.405677
	Vernon	LA	31.1083098	-93.184213
	Kent	DE	39.0861656	-75.568421

Table 3 (continued)

Species	County	State	Latitude	Longitude
T. sulcifrons	Alachua	FL	29.6747516	-82.357725
	Hamilton	FL	30.4963878	-82.947935
	Leon	FL	30.4580428	-84.277892
	Liberty	FL	30.2413745	-84.882899
	Marion	FL	29.2102021	-82.056657
	Henry	GA	33.4529985	-84.154199
	New Kent	VA	37.5051438	-76.997121
	Mecklenburg	NC	35.2464188	-80.832624
	Barbour	AL	31.8695828	-85.393209
	Talladega	AL	33.3800811	-86.165886
	Calhoun	SC	33.6748817	-80.780297
	Clarendon	SC	33.6657933	-80.216418
	Florence	SC	34.0243931	-79.702807
	Kershaw	SC	34.3387673	-80.590231
	Anderson	TN	36.1184526	-84.198459
	Cumberland	TN	35.9503754	-84.99837
	Gibson	TN	35.9966083	-88.932617
	Greene	TN	36.1753442	-82.845818
	Roane	TN	35.8478586	-84.52324
	Sevier	TN	35.7846337	-83.524182
	Colbert	AL	34.7004702	-87.804928
	Macon	AL	32.385959	-85.692653
	Lincoln	WV	38.1753494	-82.070392

Table 3 (continued)

Species	County	State	Latitude	Longitude
T. sulcifrons	York	VA	37.2431136	-76.563528
	Rowan	NC	35.6394757	-80.524787
	Georgetown	SC	33.4342485	-79.3324
	Oconee	SC	34.7534713	-83.065834
	Bullock	AL	32.1005283	-85.715682
	Chatham	NC	35.7025746	-79.255295
	Maury	TN	35.6169381	-87.077022
	Shelby	TN	35.183987	-89.895547
	Fulton	TN	33.7902745	-84.466996
	Newberry	SC	34.2898129	-81.60013
	Blount	TN	35.6872287	-83.925527
	Knox	TN	35.9932187	-83.937093
	Lexington	SC	33.9023228	-81.272201
	Whitfield	GA	34.8056101	-84.967208
	Halifax	NC	36.2574456	-77.65171
	Johnston	NC	35.5178238	-78.365709
	Lauderdale	AL	34.9014067	-87.65401
	Yell	AR	35.0026033	-93.411239
	Dougherty	GA	31.5334595	-84.216367
	Echols	GA	30.7100504	-82.893961
	Walton	GA	33.7815577	-83.73387
	Duplin	NC	34.9365355	-77.933007
	Bleckley	GA	32.4344266	-83.327853

Table 3 (continued)

Species	County	State	Latitude	Longitude
T. sulcifrons	Perry	MS	31.1720411	-88.992361
	Granville	NC	36.3040483	-78.652729
	Guilford	NC	36.0794734	-79.788907
	Hampshire	WV	39.3170745	-78.614114
	Union	TN	36.2878722	-83.837528
	Limestone	AL	34.810099	-86.981401
	Cleburne	AL	33.6745132	-85.518809
	Washington	AR	35.9790612	-94.215577
	Perry	AL	32.6384621	-87.294407
	Clarke	GA	33.9511727	-83.367335
	Tallapoosa	AL	32.8623775	-85.797498
	Spalding	GA	33.2608782	-84.284096
	Hocking	OH	39.4970642	-82.479259
	Pope	IL	37.4126944	-88.561524
	McCracken	KY	37.0539582	-88.712654
	Ripley	MO	36.6527899	-90.863866
	Chowan	NC	36.1508362	-76.607896
	Gates	NC	36.4449064	-76.700467
	Richland	SC	34.02182	-80.903053
	DeKalb	GA	33.7715442	-84.226424
	Fayette	TN	35.1971038	-89.414368
	Goochland	VA	37.7220651	-77.916525
	Pitt	NC	35.5932965	-77.374496

Table 3 (continued)

Species	County	State	Latitude	Longitude
T. sulcifrons	Johnson	IL	37.4596293	-88.880926
	Oktibbeha	MS	33.4249638	-88.879333
	Lee	AL	32.6011459	-85.355471
	Lumpkin	GA	34.5721878	-84.00267
	Wake	NC	35.7902531	-78.650312
	Charleston	SC	32.8346027	-79.95313
	Marion	AL	34.1365577	-87.887133
	Durham	NC	36.0360315	-78.876619
	Perquimans	NC	36.2058476	-76.441143

and vegetation cover variables (mean annual % cloud cover, mean annual % tree cover, mean annual % non-tree vegetation, mean annual % non-vegetated) were calculated from 2000 to 2016 from USGS datasets for a total of twelve predictor variables (Dimicili et al. 2015, Wang et al. 2016). Environmental variables were properly formatted with raster calculator tools in the SDM toolbox, averaged annually, and scaled to 1km resolution (Brown 2014). Highly correlated variables ($n \geq 0.7$) were separated with the remove highly correlated variables function in the SDM toolbox ver 2.0 (Brown 2014). Uncorrelated variables selected for models are located on Table 4.

Environmental niche modeling experiments

For visualization and mapping of distribution models, ESRI base maps and United States layers were downloaded and constructed in ArcMap 10.6 (Environmental Systems Research Institute, Redlands, CA). Training areas were determined in ArcMap with the minimum convex polygon tool with geometry type set to convex hull around the collection sites of species. Extrapolation range was set to the EPA ecoregion lvl 1 Eastern Temperate Forests (Omernik and Griffith 2014). MaxEnt software v3.4.1 was used to fit models with GPS data points of collected *Tabanus* flies and environmental data at a 1km resolution (Philips et. al 2006). Main option parameters for Maxent modeling were set to logistic output, a random seed was set to select training points randomly, and all other options remained at default. Random seed selected random samples from species presence localities for each model. Along with likely suitability ranges, MaxEnt calculated the contribution of predictors to the accuracy of the models. Twelve environmental variables were considered to include in modeling. The variables are described and identified for their significance to horse fly biology; mean annual temperature, continentality, mean annual relative humidity, mean annual precipitation, mean summer precipitation, summer heat moisture index, degree days above 18ºC, number of frost-free days mean annual

Table 4. Uncorrelated predictor variables. Uncorrelated climate and land cover predictor variables used in environmental niche modeling. Variables are listed as present or absent in the respective species distribution models they were selected for. Continentality is described as the mean temperature of the warmest month – the mean temperature of the coldest month.

Variable	***Tabanus* species**					
	T. fulvulus	*T. lineola*	*T. subsimilis*	*T. quinquevittatus*	*T. sparus milleri*	*T. sulcifrons*
Climate Variables PRISM Climate Group (http://www.prism.oregonstate.edu/)						
Mean annual temperature	No	No	No	No	No	No
Continentality	Yes	No	No	Yes	Yes	Yes
Mean annual relative humidity	Yes	Yes	Yes	Yes	Yes	Yes
Mean annual precipitation	Yes	Yes	Yes	Yes	Yes	Yes
Mean summer precipitation	Yes	Yes	Yes	No	No	No
Summer heat moisture index	No	No	No	No	No	No
Degree days above 18ºC	Yes	Yes	Yes	Yes	Yes	Yes

Table 4 (continued)

Variable	***Tabanus* species**					
	T. fulvulus	*T. lineola*	*T. subsimilis*	*T. quinquevittatus*	*T. sparus milleri*	*T. sulcifrons*
Number of frost-free days	No	No	No	No	No	No
Land Use / Land Cover MODIS (https://modis.gsfc.nasa.gov/)						
Mean annual percent cloud cover	Yes	Yes	Yes	Yes	Yes	Yes
Mean annual percent tree cover	Yes	No	No	Yes	Yes	Yes
Mean annual percent non-vegetation	No	No	No	No	No	No
Mean annual percent non-vegetated	No	Yes	Yes	Yes	Yes	Yes
Total Variables	7	6	6	7	7	7

percentage cloud cover, mean annual percentage tree cover, mean annual percentage non-tree vegetation, and mean annual percentage non-vegetated. Each variable was chosen in relation to insect development time, climate, and habitat type. All the listed environmental variables were used in county centroid modeling (Figure 11), but the variables were then tested for correlation using analysis in the SDM toolbox 2.0 to be used in the GPS-data models (Figure 12) (Brown 2014). Prediction outputs were converted to binary absence presence outputs by setting the raster threshold to greater than the minimum training presence for each species for county and GPS-data ENMs. For model analysis and validation, the training area under curve (AUC) was calculated for each species distribution model (Table 5).

Results

After testing the variables for correlation, of the 12 potential variables, six variables were used in the *T. lineola* and *T. similis* model and seven variables were used in the *T. fulvulus, T. quinquevittatus, T. sparus milleri,* and *T. sulcifrons* model listed in Table 4. *Tabanus lineola* and *T. similis* used the same six variables, while *T. quinquevittatus, T. sparus milleri,* and *T. sulcifrons* were modeled with the same seven variables. The *T. fulvulus* model used a different group of seven uncorrelated variables.

Variable contribution to model accuracy gain

Highly predictive environmental variables were identified in the GPS-data models (Figure 12) that contributed highly to model accuracy gain (Table 5). Relative humidity contributed highly to all models and continentality contributed highly to all models except for *T. lineola* and *T. subsimilis*, from which it was removed due to high correlation with other variables. The least important variable was mean annual precipitation among all species models. *Tabanus lineola* and *T. subsimilis* models used the same 6 environmental variables and produced unique distributions with slightly different variables and contribution values (Table 4). *Tabanus*

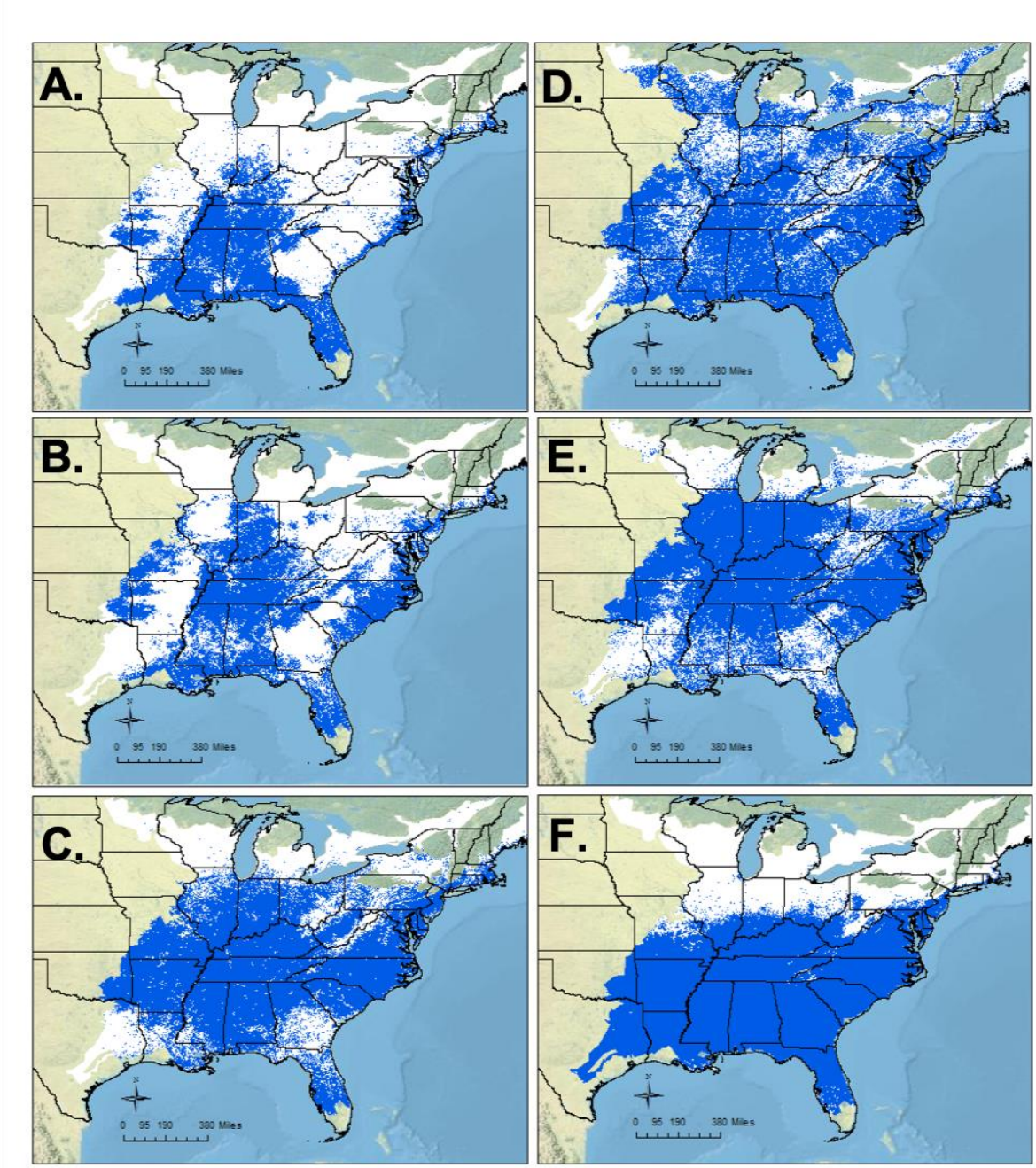

Figure 11. County centroid ENM. Minimum training presence threshold (binary) environmental niche model of the six pervasive horse flies (A) *T. fulvulus,* (B) *T. lineola,* (C) *T. subsimilis,* (D) *T. quinquevittatus,* (E) *T. sparus milleri*, and (F) *T. sulcifrons* determined with county centroid data points. White areas correspond to no predicted occurrence. Blue areas correspond to predicted occurrence.

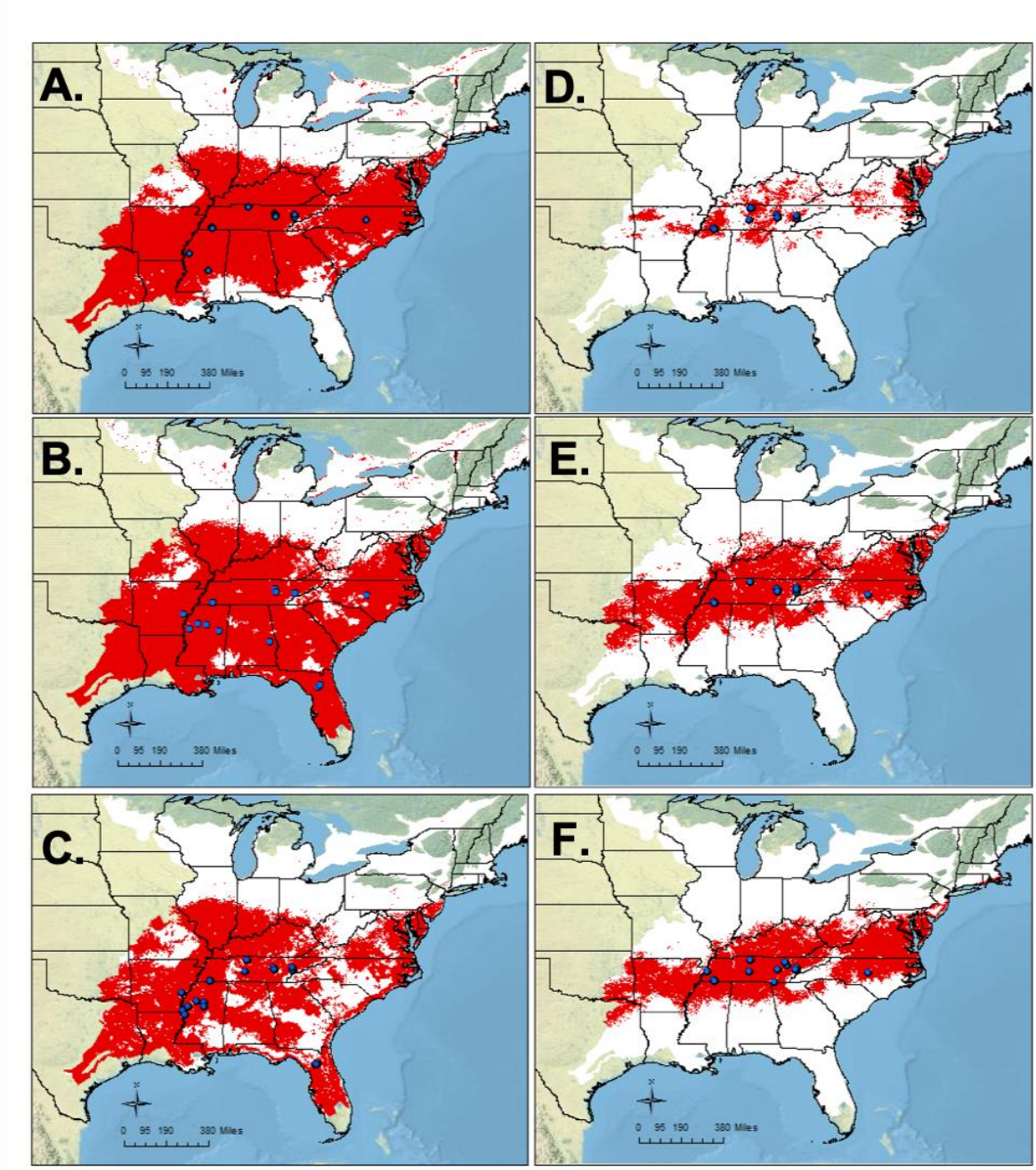

Figure 12. GPS-data ENM. Minimum training presence threshold (binary) environmental niche models of the six pervasive horse flies (A) *T. fulvulus,* (B) *T. lineola,* (C) *T. subsimilis,* (D) *T. quinquevittatus,* (E) *T. sparus milleri*, and (F) *T. sulcifrons* determined with GPS-data points (blue circles) associated with the collection sites. White areas correspond to no predicted occurrence. Red areas correspond to predicted presence.

Table 5. Variable contribution and model performance. Contribution of environmental predictor variables to accuracy of each model. AUC (area under the curve) values and omission error are also listed for each Maxent produced model.

Species	Variable	Percent Contribution	AUC		Omission Error	
			GPS	County	GPS	County
T. fulvulus	Mean annual relative humidity	43.20%	0.883	0.858	0	0
	Continentality	37.70%				
	Degree days above 18ºC	13.60%				
T. lineola	Mean annual relative humidity	66.40%	0.864	0.839	0	0
	Mean annual percent non-vegetated	37.30%				
	Mean annual percent cloud cover	6.20%				
T. subsimilis	Mean annual relative humidity	58.20%	0.914	0.823	0	0
	Degree days above 18ºC	18.20%				
	Mean annual percent non-vegetated	11.10%				

Table 5 (continued)

Species	Variable	Percent Contribution	AUC		Omission Error	
			GPS	County	GPS	County
T. quinquevittatus	Mean annual relative humidity	50.20%	0.963	0.82	0	0
	Continentality	25.40%				
	Annual average percent	11.40%				
	non-vegetated					
T. sparus milleri	Mean annual relative humidity	49.60%	0.922	0.823	0	0
	Continentality	30.20%				
	Degree days above 18ºC	9.30%				
T. sulcifrons	Continentality	54.50%	0.914	0.775	0	0
	Mean annual relative humidity	25.40%				
	Mean annual percent tree cover	8.20%				

quinquevittatus, T. sparus milleri, and *T. sulcifrons* models also used the same 7 environmental variables and produced different contribution results, with relative humidity and continentality contributing highly. Relative humidity contributed highly to the model variation in all six horse fly species. Continentality (∞C) contributed highly in *T. fulvulus, T. sulcifrons, T. sparus milleri, T. quinquevittatus* models and mean annual non-vegetated area contributed highly in *T. lineola* and *T. subsimilis* model accuracy gain.

ENM models

AUC values of GPS-data and county centroid models are 0.883 and 0.858 for *T. fulvulus,* 0.864 and 0.839 for *T. lineola,* 0.914 and 0.823 for *T. subsimilis,* 0.963 and 0.820 for *T. quinquevittatus,* 0.922 and 0.823 for *T. sparus milleri* and 0.914 and 0.775 for *T. sulcifrons* respectively. The modeled fly species have partially overlapping ranges across the state of Tennessee into surrounding states of Arkansas, Louisiana, Missouri, Illinois, Indiana, Kentucky, Virginia, North Carolina and intermittent areas in Mississippi, Alabama, Georgia, Florida, South Carolina, Ohio, West Virginia, Maryland, Delaware, and New Jersey and similar niches in the GPS-data model (Figure 11). All horse fly species have relatively widespread distribution ranges in the Eastern U.S. based on archived county collection records (Figure 12).

Conclusions

The relative abundance and population size of horse fly species vary by site and disproportionately high abundance and wide distributions of species are frequently reported in trapping surveys (Miok et al. 2006, Mikuška et al. 2016). Often the most collected horse fly is representative of the most abundant host-seeking species, but there is no conclusive evidence of this. The species chosen for modeling were selected because I could obtain adequate collection records and they were collected from at least 20 unique areas with moderate to high abundance. The six modeled *Tabanus species* were identified as pervasive and can be difficult

to accurately identify with morphology. Here I was able to define clearly the different environmental predictors associated with the pervasive horse flies, which correspond to their habitat and ecology.

Similarities in the suitability ranges of the six horse flies modeled occur across much of the East Central and South Atlantic U.S. Dissimilarities are noted in the narrower suitability ranges of *T. quinquevittatus* and *T. sparus milleri* to the other more wide-ranging models. *Tabanus quinquevittatus* is the most restricted in predicted range based on GPS-data records, but very widely distributed based on county records. The distribution of *T. sparus milleri* is similar to the distribution of T. sulcifrons. *Tabanus lineola* and *Tabanus subsimilis* appear to be the most widely distributed species, likely occuring across most of the Eastern U.S.

High relative humidity contributed highly to all GPS-data models and therefore is likely an important abiotic factor in the occurrence of the species modeled. This is likely because humid environments provide adequate immature and adult horse fly habitats. Whether this is realized in their fundamental niche as a result of vegetation type or other habitat requirements is yet to be determined. In the GPS-data models, Louisiana, Georgia, South Carolina, Kentucky, Virginia, Maryland, and Delaware did not have occurrence records recorded from them to fit and train the model however, suitability was predicted in these states for the six species modeled. These areas likely correspond highly with the fundamental niche of each horse fly species and have suitable habitat based on the models. *Tabanus fulvulus* is known from the Central and Eastern U.S. and its presence is highly correlated with high relative humidity similar to the other greenhead flies modeled (*T. lineola, T. subsimilis,* and *T. quinquevittatus*). *Tabanus lineola* is known from Eastern North America, the Caribbean, and Central America and is one of the most widespread species included in this study. *Tabanus subsimilis* is known from North and Central America and shares a similar distribution to *T. lineola*. *Tabanus quinquevittatus*, *T. sparus*

milleri, and *T. sulcifrons* are known from the Eastern U.S. and share similar distributions based on both ENMs (Teskey and Pechuman 1983). This is likely due to similar collection records and as a result of similar training areas for modeling.

Distributions of the six horse fly species correspond to suitability based on GPS-data modeling and correspond to the fundamental niche. Precise suitability ranges and niches cannot be inferred from county-centroid models. The county-centroid models offer to fill in categorical (state, region) information gaps left by the lack of representation in the GPS collection record models. We created both GPS and county-centroid distribution maps to illustrate the importance of recording GPS data at trap sites. We also recognize that the GPS-data model is incomplete with false negatives, but low in false positives. However, the county-centroid model is likely higher in false positives and lower in false negatives. Nevertheless, we find utility in both models as it is likely that the county-centroid model will yield more occurrence points which can then be interpreted with the more accurate GPS model.

Maxent modeling tends to limit commission error (false positive) however, models can produce omission error (false negative) if there are missing collection data. Areas known to have horse fly populations based on county museum records (Table 3) but do not predict suitability within this extent must be fit with more GPS collection data to corroborate current knowledge with likelihood at a precise scale (1km). This is crucial to understanding the biological underpinnings of habitat selection of each species based on abiotic and biotic interactions that are omitted with coarse collection data.

Like most models, limitations include the low number of GPS collection records and limited archived (in museum collections) horse flies with geographic coordinates. Due to collection gaps in Georgia, South Carolina and Alabama, models could not be fit with data most representative

of *Tabanus* species full distributions. Uncorrelated predictor variables used to train the models were not consistent with the uncorrelated variables in the model extrapolation range of the EPA ecoregion (eastern temperate forests). Uncorrelated variables designated for the extrapolation extent may add to model uncertainty and included mean annual relative humidity, mean annual precipitation, mean summer precipitation, summer heat moisture index, mean annual % cloud cover, mean annual % tree cover, and mean annual % non-vegetated. Also contributing to model uncertainty are horse fly dispersion patterns and behavior. Based on previous studies, we know that some horse flies can travel 1-2km away from their breeding and development sites, but the rate at which this occurs is uncertain (Cooksey and Wright 1987). With dispersal rates of *Tabanus* species inconclusive, more information is needed to apply dispersal behavior to distribution mapping.

A benefit to this study and future studies interested in utilizing distribution modeling will be the widespread adoption of GPS records for collected taxa. GPS availability to the public and researchers has only been made available relatively recently and should be considered a standard for specimens archived for future research applicability. The produced niche models will be useful in identifying pestiferous *Tabanus* and aid in future research regarding *Tabanus* management and control. GPS-data models should be confirmed for their accuracy by collecting in areas with reported high suitability. Areas of most interest to collect are those with no previous GPS records, but still displayed as suitable such as Virginia, Maryland, and Delaware. Furthermore, projecting horse fly distributions in the future with changing climatic conditions and remotely sensed data will have implications on human and animal health, ecosystem health, and food security. As the global climate warms and other climatic variables continue their current trends, *Tabanus* populations are expected to become more competitive in most of their current habitats and need to be considered threats both contemporarily and in the future.

Conclusion

A total of 18,248 horse flies were collected from 2014-2018 totaling 40 species from sites in Alabama (1 site), Arkansas (1 site), Florida (14 sites), Mississippi (14 sites), North Carolina (2 sites), and Tennessee (42 sites). Horse flies were utilized for molecular and niche analysis with the sequencing of the COI gene and environmental niche modeling.

Barcoding to aid in species identification

The COI phylogenetic tree constructed with sequences generated from 206 specimens representing 40 species was analyzed with Bayesian inference and produced 9 distinctive clades that distinguished most species with posterior probability support. Sequences that corroborated *Tabanus* species' morphological data and were recovered with monophyly included *T. nigrovittatus, T. mularis, T. quinquevittatus, T. superjumentarius*, *T. nigripes, T. trimaculatus, T. sparus milleri, T. lineola, T. subsimilis, T. pumilus, T. fairchildi, T. equalis, T. turbidus, T. rufofrater, T. americanus, T. reinwardtii, T. calens, T. imitans, T. atratus, T. aranti, T. stygius, T. nigriscens, T. cymatomorphous*, and *T. proximus*. *Tabanus species* that were occasionally recovered in morphologically separate lineages with low support included *T. petiolatus*, *T. fulvulus, T. longiusculus, T. sackeni, T. sublongus, T. longus*, *T. molestus*, *T. mixis*, *T. moderator*, *T. sulcifrons* and *T. abdominalis*. Results based on COI sequence data are inconclusive to distinguish these species within their respective closely related complexes and more research is required. Considering these species are significant pests, taxonomic resolution of the abovementioned species should be prioritized with future molecular studies.

Environmental niche modeling to aid in identification

The six horse fly species (*T. fulvulus, T. lineola, T. subsimilis, T. quinquevittatus, T. sparus milleri* and *T. sulcifrons*) modeled have partially overlapping ranges and overlapping niches. Suitable ranges of the horse fly species correspond to suitability based on GPS records and correspond to the fundamental niche. Relative humidity contributed highly to the model

accuracy in all six horse fly species. Continentality (∞C) contributed highly in *T. fulvulus, T. sulcifrons, T. sparus milleri, T. quinquevittatus* models and mean annual non-vegetated area in *T. lineola* and *T. subsimilis* models. Similarities in the suitability ranges occur across the state of Tennessee into surrounding states, Arkansas, coastal Virginia, and the peninsular region of Maryland and Delaware. Dissimilarities are noted in the more restricted ranges of *T. quinquevittatus*, *T. sparus milleri, and T. sulcifrons* to the other more wide-ranging species. *T. quinquevittatus* is the most restricted in predicted range based on GPS records, but very widely distributed based on county records. *Tabanus lineola* and *T. subsimilis* appear to be the most widely distributed species predicted to likely occur across the Central and Southeastern U.S., and *T. sulcifrons,* is distributed across the Eastern U.S.

Study implications

To improve horse fly management decisions and tactics in the future, evaluations for the development of effective integrated pest management tools such as economic injury levels are required (Pedigo and Higley 1992). The difficulty of Tabanidae control is discussed; in the 1950s widespread insecticide usage over the aquatic and semi-aquatic larval habitat did lower Tabanidae numbers in some cases, but damaged the local environment making it unsustainable and not advisable (Wall and Marganian 1973). Specific insecticides that target larval horse flies in these expansive habitats have yet to be developed. Altering water levels has been suggested and attempted on a few occasions with infeasible assumptions and inconclusive results as well as environmental drawbacks (Anderson 1969). With immature control unachievable at the moment, adult trapping is a preferred method for some measure of control (Teskey and Pechuman 1983). While collecting many horse flies may offer some relief in immediate areas (Wilson 1968), it is doubtful that trapping adult horse flies alone decreases livestock infestations over time. This is due to several aspects of horse fly biology including a high fecundity (female can lay 100-1000 eggs), and the high peaks in adult activity with

populations far too large to completely trap out (Jones and Anthony 1964). Use of a topical pyrethroid spray and insecticide-impregnated ear tags have some level of horse fly efficacy, although flies will often feed despite applications before reacting (Leprince et al. 1991). Previous studies demonstrated that ear tags have longer durations of efficacy on horse fly mortality and knockdown that can extend to 2-3 weeks (Presley and Wright 1986). Fly knockdown rates during this efficacy window have been shown to be as high as 80% for some species of horse flies, but this is typically seen within just the first few days after insecticide application and tapers off steeply over time.

A number of flies travel long-distance and come from off-pasture sites to feed on livestock. Use of barriers to prevent horse flies from entering pastures has not been thoroughly investigated but is a potentially viable control option. Two different barriers can be attempted to ward specific areas. The first is a physical net barrier, which is treated with an insecticide and erected around the pasture. Net barriers have been tested in the past and may hold promise based on the observation of horse flies flying around as opposed to over the top of screen barriers of at least 8 feet in height (Barros and Foil 2007). The second method is the use of an insecticide barrier applied to foliage/vegetation. Insecticide barriers are used to minimize mosquito and other biting fly populations and could target resting adult horse flies as well. If a livestock operator suffers a disease outbreak they suspect to be caused by horse flies, it is recommended that control measures be taken to the extent that they achieve control to no more than 10 horse flies per animal if possible (Desquesnes et al. 2009). This can be attempted with the suggestions listed above to fit the producer's needs. If only a few infected animals can be identified as symptomatic of an infectious pathogen, it is recommended to quarantine infected animals 200m away from healthy individuals (Barros and Foil 2007).

Future directions

Throughout the collection period of the study, more immature and male collections would have been beneficial if time and funding were not an issue. More male horse flies could have been collected with the use of malaise traps around immature habitats (Smith et al. 1994). Capturing *Tabanus* larvae would have allowed the testing of COI barcoding for identification of immature flies, but none were collected despite attempts. Increased geographic and species diversity of collections would also improve the study given the time and resources. Evaluations for the development of effective integrated pest management tools for Tabanidae through the measurement of parameters established by the Economic Injury Level will lead to unexplored and promising research (Pedigo and Higley 1992). Of these parameters, injury units per pest, damage per injury unit, and proportional reduction in pest attack vary widely among pest species and must be measured through experimental trials (Pedigo and Higley 1992). The ability to quantify these parameters successfully is more feasible than ever with the use of these data for identification of horse flies.

References

(Equine Infectious Anemia Distribution Maps) U.S. DEPARTMENT OF AGRICULTURE. **2017**. Equine Infectious Anemia Distribution Maps.

Anderson, J. F. 1969. The Temporary Impoundment of Salt Marshes for the Control of Coastal Deer Flies.

Axtell, R. C. 1976. Coastal horse flies and deer flies (Diptera : Tabanidae).

Baldacchino, F., L. Puech, S. Manon, L. R. Hertzog, and P. Jay-Robert. **2014**. Biting behaviour of Tabanidae on cattle in mountainous summer pastures, Pyrenees, France, and effects of weather variables. Bull. Entomol. Res. 104: 471–479.

Banerjee, D., V. Kumar, A. Maity, B. Ghosh, K. Tyagi, D. Singha, S. Kundu, B. A. Laskar, A. Naskar, and S. Rath. **2015**. Identification through DNA barcoding of Tabanidae (Diptera) vectors of surra disease in India. Acta Trop. 150: 52–58.

Barros, A. T. M., and L. D. Foil. **2007**. The influence of distance on movement of tabanids (Diptera: Tabanidae) between horses. Vet. Parasitol. 144: 380–384.

Brown, J. L. 2014. SDMtoolbox: A python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. Methods Ecol. Evol. 5: 694–700.

Brun, R., H. Hecker, and Z. R. Lun. **1998**. Trypanosoma evansi and T. equiperdum: Distribution, biology, treatment and phylogenetic relationship (a review). Vet. Parasitol. 79: 95–107.

C. Daly, G. H. Taylor, W. P. Gibson, T. W. Parzybok, G. L. Johnson, and P. A. Pasteris. **2000**. High-Quality Spatial Climate Data Sets for the United States and Beyond. Trans. ASAE. 43: 1957–1962.

Cárdenas, R. E., J. Buestán, and O. Dangles. **2009**. Diversity and distribution models

of horse flies (Diptera: Tabanidae) from Ecuador. Ann. la Soc. Entomol. Fr. 45: 511–528.

Cooksey, L. M., and R. E. Wright. **1987**. Flight Range and Dispersal Activity of the Host-seeking Horse Fly, Tabanus abactor (Diptera: Tabanidae), in North Central Oklahoma. Environ. Entomol. 16: 211–217.

Costa, F. O., and G. R. Carvalho. **2010**. New insights into molecular evolution: Prospects from the barcode of life initiative (BOLI). Theory Biosci. 129: 149–157.

Cywinska, A., M. A. Hannan, P. G. Kevan, R. E. Roughley, M. Iranpour, and F. F. Hunter. **2010**. Evaluation of DNA barcoding and identification of new haplomorphs in Canadian deerflies and horseflies. Med. Vet. Entomol. 24: 382–410.

Deng, S., and C. Hiruki. **1991**. Amplification of 16S rRNA genes from culturable and unculturable Mollicutes. J. Microbiol. Methods. 14: 53–61.

Desquesnes, M., F. Biteau-Coroller, J. Bouyer, M. L. Dia, and L. Foil. **2009**. Development of a mathematical model for mechanical transmission of trypanosomes and other pathogens of cattle transmitted by tabanids. Int. J. Parasitol. 39: 333–346.

Desquesnes, M., A. Dargantes, D. H. Lai, Z. R. Lun, P. Holzmuller, and S. Jittapalapong. **2013**. Trypanosoma evansi and surra: A review and perspectives on transmission, epidemiology and control, impact, and zoonotic aspects. Biomed Res. Int. 2013.

Drees, B., L. Butler, and L. L. Pechuman. **1980**. Horse flies and deer flies of West Virginia : an illustrated key (Diptera, Tabanidae). West Virginia Univ. Libr. 674.

Drummond, A. J., W. Xie, and J. Heled. **2012**. Bayesian Inference of Species Trees

from Multilpcus Data using *BEAST. 1–18.

Ferrier, S., A. Guisan, J. Elith, C. H. Graham, R. P. Anderson, M. Dudı, R. J. Hijmans, F. Huettmann, J. R. Leathwick, A. Lehmann, J. Li, L. G. Lohmann, B. A. Loiselle, G. Manion, C. Moritz, M. Nakamura, Y. Nakazawa, J. M. Overton, A. T. Peterson, S. J. Phillips, K. Richardson, R. Scachetti-pereira, R. E. Schapire, S. Williams, M. S. Wisz, and N. E. Zimmermann. **2006**. Novel methods improve prediction of species distributions from occurrence data. 2.

Foil, L. D. 1989. Tabanids as vectors of disease agents. Parasitol. Today. 5: 88–96.

Foil, L. D., and J. a Hogsette. **1994**. Biology and control of tabanids, stable flies and horn flies. Rev. Sci. Tech. 13: 1125–1158.

Folmer, O., M. BLACK, W. HOEH, R. Lutz, and R. Vrijenhoek. **1994**. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotechnol. 3: 294–299.

Friend, W. G., and J. G. Stoffolano. **1991**. Feeding-Behavior of the Horsefly Tabanus-Nigrovittatus (Diptera, Tabanidae) - Effects of Dissolved Solids on Ingestion and Destination of Sucrose or Atp Diets. Physiol. Entomol. 16: 35–45.

Goodwin, J. T., and D. M. Bastiaan. **1996**. The Horse and Feer Flies of Texas. Southwest. Entomol.

Goodwin, J. T., B. A. Mullens, and R. R. Gerhardt. **1985**. The Tabanidae of Tennessee. Knoxville, Tennessee.

Hall, T. A. 1999. Bioedit: A User-friendly biological sequence alignment editor and analysis program. Nucl. Acids Symp. 41: 95–98.

Hawkins, J. A., W. V Adams, B. H. Wilson, C. J. Issel, and E. E. Roth. **1976**.

Transmission of equine infectious anemia virus by Tabanus fuscicostatus. J. Am. Vet. Med. Assoc. 168: 63–4.

Hawkins, J. A., J. N. Love, and R. J. Hidalgo. **1982**. Mechanical transmission of anaplasmosis by tabanids (Diptera: Tabanidae). Am. J. Vet. Res. 43: 732–734.

Hebert, P. D. N., S. Ratnasingham, and J. R. de Waard. **2003**. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proc. R. Soc. B Biol. Sci. 270: S96–S99.

Jones, C. M., and D. W. Anthony. **1964**. The Tabanidae of Florida. Agric. Res. Serv. USDA.

Krinsky, W. L. 1976. Animal disease agents transmitted by horse flies and deer flies (Diptera: Tabanidae). J. Med. Entomol. 13: 225–275.

Kristen Bartlett, Steven R. ALM, Roger Lebrun, H. G. 2002. The Horse and Deer flies (Diptera, Tabanidae) of Rhode Island. Entomol. Soc. Am. 2: 144.

Kunz, S. E., K. D. Murrell, G. Lambert, L. F. James, and C. E. Terrill. **1991**. Estimated losses of livestock to pests. CRC Handb. Pest Mangement Agric. vol 1. I: 68–69.

Leprince, D. J., L. D. Foil, and R. L. Byford. **1991**. Evaluation of Pyrethroid Ear Tag and Spray Treatment of Cattle Against Horse Flies (Diptera Tabanidae). Louisiana Exp. Stn.

Lessard, B. D., and D. K. Yeates. **2011**. New species of the Australian horse fly subgenus Scaptia (Plinthina) Walker 1850 (Diptera: Tabanidae), including species descriptions and a revised key. Aust. J. Entomol. 50: 241–252.

Livestock, I. 1977. B. Gill, Trypanosomes and Trypanosomiases of Indian Livestock,

Edited by ICAR, Indian Council of Agricultural Research.

Meier, R., K. Shiyang, G. Vaidya, and P. K. L. Ng. **2006**. DNA barcoding and taxonomy in diptera: A tale of high intraspecific variability and low identification success. Syst. Biol. 55: 715–728.

Mikuška, A., S. Mlinarić, L. Begović, and E. Curran. **2016**. Comparative efficiency of traps for horse fly (Diptera: Tabanidae) survey in riparian oak-ash forests in Danube floodplain. Eur. J. Entomol. 113: 531–536.

Mooring, M. S., D. T. Blumstein, D. D. Reisig, E. R. Osborne, and J. M. Niemeyer. **2007**. Insect-repelling behaviour in bovids: Role of mass, tail length, and group size. Biol. J. Linn. Soc. 91: 383–392.

Morita, S. I., K. M. Bayless, D. K. Yeates, and B. M. Wiegmann. **2016**. Molecular phylogeny of the horse flies: A framework for renewing tabanid taxonomy. Syst. Entomol. 41: 56–72.

N.S. Krishna Rao, S. M. 1958. Tabanus Flies as Transmitters of Anthrax - A Field Experience. Indian Vet. J. 38.

Nalen, C. M. Z., D. L. Kline, B. D. Sutton, G. Müller, and J. E. Cilek. **2015**. An Annotated Checklist of the Horse Flies, Deer Flies, and Yellow Flies (Diptera: Tabanidae) of Florida. Florida Entomol. 98: 479–488.

Nascimento, F. F., M. Dos Reis, and Z. Yang. **2017**. A biologist's guide to Bayesian phylogenetic analysis. Nat. Ecol. Evol. 1: 1446–1454.

Omernik, J. M., and G. E. Griffith. **2014**. Ecoregions of the Conterminous United States: Evolution of a Hierarchical Spatial Framework. Environ. Manage. 54: 1249–1266.

Pearson, R. G., J. R. Raxworthy, M. Nakamura, and A. T. Peterson. **2007**. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. J. Chem. Technol. Metall. 52: 277–287.

Pedigo, L. P., and L. G. Higley. **1992**. The Economic Injury Level Concept and Environmental Quality: A New Perspective. Am. Entomol. 38: 12–21.

Pérez, M. P., J. Palacio, M. P. Santolaria, M. D. C. Aceña, G. Chacón, M. T. Verde, J. H. Calvo, M. P. Zaragoza, M. Gascón, and S. García-Belenguer. **2002**. Influence of lairage time on some welfare and meat quality parameters in pigs. Vet. Res. 33: 239–250.

Perich, M. J., R. E. Wright, and K. S. Lusby. **1986**. Impact of Horse Flies (Diptera: Tabanidae) on Beef Cattle. J. Econ. Entomol.

Philip S. Corbet, A. J. H. (East A. V. R. I. 1962. Diptera Swarming High Above the Forest Canopy in Uganda With Special Reference to Tabanidae.

Phillips, S. B., V. P. Aneja, D. Kang, and S. P. Arya. **2006**. Maximum entropy modeiling of species geographic distributions. Int. J. Glob. Environ. Issues. 6: 231–252.

Presley, S. M., and R. E. Wright. **1986**. Field Test of Pyrethroid Ear Tags, Sprays, and a Pour-on Formula- Tion for Control of Horse Flies on Cattle. October. 369–373.

Quercia, O., F. Emiliani, F. G. Foschi, and G. F. Stefanini. **2008**. The wasp-horsefly syndrome. Med. Interna Ital. 61–63.

Raxworthy, C. J., C. M. Ingram, N. Rabibisoa, and R. G. Pearson. **2007**. Applications of ecological niche modeling for species delimitation: A review and empirical evaluation using day geckos (Phelsuma) from Madagascar. Syst. Biol. 56: 907–

923.

S. Miok, D.A. Carlson, E.S. Krafsur, L. D. F. 2006. Performance of the NZI and other traps for biting flies in North America. Bull. Entomol. Res.

Smith, S. M., D. A. Turnbull, and P. D. Taylor. **1994**. Assembly, mating, and energetics of Hybomitra arpadi (Diptera: Tabanidae) at Churchill, Manitoba. J. Insect Behav. 7: 355–383.

Suchard, M. A., P. Lemey, G. Baele, D. L. Ayres, A. J. Drummond, and A. Rambaut. **2018**. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus Evol. 4: 1–5.

Teskey, H. 1969. LARVAE AND PUPAE OF SOME EASTERN NORTH AMERICAN TABANIDAE (DIPTERA). Mem. Entomol. Soc. Canada, 101(S63), 5-147.

Teskey, H. J., and L. L. Pechuman. **1983**. The Diptera, or True Flies, of Illinois I. Tabanidae. 33.

Tidwell, M. A., W. D. Dean, G. P. Combs, D. W. Anderson, W. O. Cowart, and R. C. Axtell. **1972**. Transmission of hog cholera virus by horseflies (Tabanidae: Diptera). Am. J. Vet. Res. 33: 615–622.

Wall, W. J., and V. M. Marganian. **1973**. Control of Salt Marsh Culicoides and Tabanus Larvae in Small Plots with Granular Organophosphorus Pesticides and The Direct Effect on Other Fauna.

Wilson, B. H. 1968. Reduction of Tabanid Populations on Cattle with Sticky Traps Baited with Dry Ice. J. Econ. Entomol. 1–4.

Vita

Travis is a native of Gainesville Florida where he attended the University of Florida and received a bachelor's degree in biology. While exploring research interests, opportunities with the United States Department of Agriculture in the labs of Dr. Dewayne Shoemaker and other research entomologists led him to pursue these interests after graduating.